Representation of data analysis results in multidimensional parameter space

Sergey V. Zykin
Sobolev Institute of Mathematics, Siberian Branch of the RAS, Acad. Koptyug avenue, h. 4, Novosibirsk, Russian Federation, 630090
E-mail: svzykin01@gmail.com

Abstract. In data analysis, presenting the boundaries between the classes of objects is considered a minor issue in most cases. However, the subsequent use of the analysis results (for example, in diagnostic tasks or in the acquisition of the necessary object properties by controlling the parameters) should be based on the boundaries delineation and the accuracy of their description. This underlines the need to develop and use universal methods for presenting the data analysis results. This paper considers the data model for applied decision support systems, in which one of the components is graphic data, i.e. domains in multidimensional space, bounded by general surfaces. A mathematical model is proposed, which outlines the range of possible graphic applications rather strictly. The paper proposes a meaningful approach to describing the boundaries delineation error. The material considered can serve as a basis for technology of data analysis results storage and use.

1. Introduction
Nowadays the methods of data analysis are widely used to detect patterns in the parameter space of the objects under investigation. As a result of the analysis, objects can be distributed among many different classes. In cluster analysis methods, round-shaped domains are most commonly used, and diagnostics is performed through the distance from centroids. If there are mutual nonlinear dependencies of the parameters, this approach becomes inappropriate, since the centroid may be beyond the boundaries of the corresponding parameter space domain. More precisely-defined domain boundaries are assumed in the discriminant analysis. In this case, piecewise linear functions are used for nonlinear dependencies approximation, this significantly complicating the preparation of the initial data due to the connectivity determination in multidimensional space. It also introduces additional error in the boundaries detection.

The use of data analysis results underlies the information support of the decision making process. Such information must be subjected to examination and approval before being used. However, this situation doubles the cost of data analysis, and it is not always done in a qualified manner. The way out of this situation is to organize the storage of information, which has passed a qualified examination, in a unified form in a database.

Let us consider an example of graphical data in the parameter space $P_1$ and $P_2$ (Fig. 1), for which a mathematical model is proposed in this work.

Such representation of the data analysis results corresponds to the graphical presentation of domains in the parameter space. Each area corresponds to one class of the objects of the same type.
The way the boundaries and their errors are determined depends on the objectives set and the means of presenting source data. Publication [1] develops the method of Lan and DeMets who proposed a spending function approximating the O’Brien-Fleming boundary based on the Brownian motion process. The method was extended to a family with additive boundaries and conditional error.

In [2] a new method is proposed to evaluate uncertainty on the boundaries of spatial clusters identified through the spatial scan statistic. For each spatial data location, the value of belonging to the true cluster is calculated. Through simulation modelling, it is shown that that method provides a way to define and visualize the certainty or uncertainty of each location. The method has been implemented in boundary analysis for the circular spatial scan statistic of aggregated data.

Procedures for partitioning large collections of highly intermixed datasets of different classes are given in [3]. Hyper-spherical or hyper-ellipsoidal clusters are used as cluster boundaries. The incremental procedures generate a minimal number of such clusters. Each cluster, however, is to contain a maximal number of data points of the same class. The developed procedures present an extension of the move-to-front algorithms and can be applied to cluster boundaries modeling to solve numerous data analysis problems.

Many papers do not discuss the problem of cluster delineation and their measurement error. For example, the peculiarities of clustering the ecological data monitoring are discussed in [4]. A complex assessment of the environmental situation is proposed, that helps to identify the structural links between the monitoring indicators. However, “pixel” cluster views are used in the paper. One of the drawbacks of such cluster view is descriptive redundancy and poor scalability.

This article proposes a data model for cluster description and delineation in the space of parameters obtained through data analysis.

2. Data model requirements
Mathematical description of cluster boundaries in the parameter space should meet the following requirements.

- Element with dimension $i$, $i = 1, n$, is bounded by elements with dimension $i - 1$, $i - 2$, ..., 0, simultaneously belonging to the source element and the adjacent elements with the same dimension. To implement this principle, the element numbers have been introduced and for each element with dimension $i - 1$, the numbers of elements with dimension $i$ to which it belongs are specified.
- There are some points which the boundary must go through precisely.
- There are some points for which the boundary passes within the measurement error of these points.

![Figure 1. Example of cluster delineation](image)
• There are some points for which the boundary passes at the maximum distance.
• The boundaries must meet the topological requirements of the application field. Elements with dimension \( n - 1, n - 2, \ldots, 1 \) must contain no false extrema and inflection points. This may be due to data errors or an underestimated discrepancy (error) value. In the first case, it is necessary to adjust the data. In the second case, it is necessary to increase the discrepancy, i.e. to match it with the real accuracy of available data.
• The junction of elements (surfaces) to be approximated separately is carried out in such a manner that their continuation belongs to the element with a larger dimension for which the source elements are not boundary. This problem usually arises in case of lacking experimental data and is solved by adding new points obtained with the calculation method.

Various tasks can be solved using the representation of data analysis results, for example:
• object recognition by a set of parameters;
• object state determination;
• controlling the transition of the object to the target state;
• minimizing energy costs during object management, etc.

Unifying the representation of data analysis results will create shared information resources. As the data in such system are structured, it is reasonable to use database technologies for their storage and processing.

3. Mathematical model of data view
Growing interest to the ways of graphic information representation and processing by means of the primitives has been observed for many years [5, 6]. On the one hand, it is caused by the necessity to create automated systems utilizing graphic information; on the other hand, it is due to increased technical capabilities of computing facilities.

In general, the results of data analysis are \( n \)-dimensional objects in parameter space. The manner in which data are presented should ensure efficient data storage and processing. It is necessary to mention that the mathematical model considered further can be used to describe a wide class of real objects as its properties are of a rather general character.

3.1. Selection of data representation method
There are many ways to represent real objects graphically. As a rule, they are used to describe two-dimensional and three-dimensional objects. The following main factors influence the choice of this or that method:
• geometrical properties of the objects;
• applied tasks solved in this information system.

By geometrical properties one means structure (topology) of the objects and required accuracy (adequacy) of their description.

Pixel representation of the objects [4, 7] is used to process various types of photographic materials and allows practically any analysis tasks to be solved. However, due to low accuracy, this method is not used in areas where a clear delineation of objects in the graphics is required. For example, it cannot be used in computer-aided design, surveying, etc., that is wherever scaling at large ranges is required.

For cluster boundaries it is reasonable to use approximations of functional dependencies. This is connected not only with the need for scaling, although it is relevant for diagnostics near the boundaries, but also for building the boundaries themselves in discriminant analysis. In [8] it is proposed to use parameter functions to represent multidimensional graphical information:

\[
y = y(t), \quad x = x(t), \quad n = 2.
\] (1)
Such a method has an undeniable advantage when there is no need for single-valued calculation of the function \( y = y(x) \) while calculating some characteristics, for example, obtaining two-dimensional sections, projections, volume calculation, etc. However, these algorithms are not enough when working with data analysis results. For further processing of information obtained by approximating the cluster boundaries, it is important to calculate the function \( y = y(x) \), so preference is given to single-valued functions. In this case, many-valued surfaces are broken down into single-valued domains and approximated separately.

It should also be noted that the parametric representation of surfaces at \( n \geq 3 \) requires the introduction of connectivity [9] for the points of the assumed surfaces, which does not guarantee the correct boundary formation and complicates the preparation of the source data.

### 3.2. Formal description of the data model

The results of data analysis should be represented by domains in the source parameter space. The space domains (clusters) are separated by the surfaces having one less dimensions than the areas themselves. The surfaces form a ”framework” of data view, bounded in the space of parameter changes.

Let \( E^n \) be Cartesian space of dimension \( n \) with orthogonal basis \( e_1, e_2, \ldots, e_n \). The scalar product and the norm of \( E^n \) are determined in the usual way:

\[
(a, b) = \sum_{i=1}^{n} a_i b_i, \quad ||a|| = (a, a),
\]

where \( a_i, b_i \) are real numbers, \( i = 1, \ldots, n \).

In space \( E^n \) let us consider the domains \( \Omega_i, i = 1, \ldots, N \), where \( N \) is finite.

**Property 1.**

\[
\bigcup_{i=1}^{N} \Omega_i = E^n, \quad \Omega_i \neq \emptyset.
\]

**Property 2.**

\[
\Omega_i \cap \Omega_j = \Gamma_{ij}, \quad i, j = 1, \ldots, N, i \neq j,
\]

where \( \Gamma_{ij} \) is closed, totally bounded domain in \( E^n \) of dimension \( n-1, n-2, \ldots, 1, 0 \). There are pairs of indices \((i, j)\), for which \( \Gamma_{ij} = \emptyset \). Moreover, \( \Gamma_{ij} \) are such that any closed connected set having a non-empty intersection with more than one domain \( \Omega_i \) will definitely have a non-empty intersection with some boundary \( \Gamma_{ij} \).

Each \( \Gamma_{ij} \neq \emptyset \) is set with a certain error \( \Delta_{ij} \), where \( \Delta_{ij} \) is a connected \( n \)-dimensional closed totally bounded domain in \( E^n \) and \( \Gamma_{ij} \subset \Delta_{ij} \).

Let us consider the projection operator \( W_{l_1} \) along the coordinate \( x_{l_1} \). \( W_{l_1} \) is a matrix of \( n \times n \) with ones on the main diagonal except for the line \( l_1 \). All other elements of the matrix are zero. It is not difficult to show that \( W_{l_1} \) calculates vector \( y \in E_{l_1}^{n-1} \) for each vector \( x \in E^n \) where \( E_{l_1}^{n-1} \) is a subspace in \( E^n \), consisting of the vectors \((x_1, x_2, \ldots, x_{l_1-1}, 0, x_{l_1+1}, \ldots, x_n)\). Besides, \( y = W_{l_1} x \) corresponds to the minimum of the norm \( ||x - y|| \) for any \( y \in E_{l_1}^{n-1} \) and the difference of vectors \( x - y \) is orthogonal to \( E_{l_1}^{n-1} \).

**Property 3.** For all \( \Gamma_{ij} \) with the dimension \( n-1 \) there is such index \( l_1 \) that

\[
\Gamma_{ij} = \{ x \in E^n \mid x_{l_1} = A_{ij}^{q_1} (x'), \quad x' \in \Omega_{ij}^{q_1} = W_{l_1} \Gamma_{ij}^{q_1} \}
\]

\[
\Gamma_{ij} = \bigcup_{q_1} \Gamma_{ij}^{q_1},
\]
where \( k(i,j) \) are finite positive integers; \( \Gamma_{ij}^{q_1} \) are closed domains in \( E^n \), the definition of \( \Gamma_{ij} \) suggests that all \( \Gamma_{ij}^{q_1} \) are totally bounded; \( \Omega_{ij}^{q_1} \) is a domain in \( E^n \), from the definition of \( \Gamma_{ij} \) and \( W_l \) it follows that \( \Omega_{ij}^{q_1} \) are closed totally bounded sets; \( A_{ij}^{q_1} \) is a single-valued and continuous function in \( \Omega_{ij}^{q_1} \).

Therefore, \( A_{ij}^{q_1}, q_1 = 1, k(i,j) \), together set the boundary \( \Gamma_{ij} \) for the partition of the domains \( \Omega_i \) and \( \Omega_j \). Since \( A_{ij}^{q_1} \) is single-valued, the indices \( i \) and \( j \) can be defined more precisely as follows.

Let the lower first index \( i \) be the number of the domain corresponding to the larger values of \( x_{l_1} \) relative to \( \Gamma_{ij} \), and the lower second index \( j \) be the number of the domain corresponding to the smaller values of \( x_{l_1} \). Each boundary \( \Gamma_{ij} \) will be matched with the triple \((i, j, q_1)\).

Through the given point \( x^0 \in E^n \) the following sets will be defined

\[
S = \{ x \in E_n | W_l x = W_l x^0 \},
\]

\[
S(i_1) = \bigcup_{(i,j,q_1)} (\Gamma_{i,j}^{q_1} \bigcup \Gamma_{i,j}^{q_2}).
\]

**Property 4.** The sets \( S \cap S(i_1) \cap S(i_2), i_1 \neq i_2 \) must have no more than one point for any \( x^0 \in E^n \).

**Property 5.** Let us consider the boundaries \( \Gamma_{ij}^{q_1} \) of the domains \( \Omega_{ij}^{q_1} \subset E^{n-1} \). It will be assumed that for all \( \Gamma_{ij}^{q_1} \) there is such index \( l_2 \) that

\[
\Gamma_{ijp}^{q_1,q_2} = \{ x' \in E^{n-1} | x_{l_2} = A_{ijp}^{q_1,q_2}(x''), \ x'' \in \Omega_{ij}^{q_1} = W_{l_2} \Gamma_{ijp}^{q_1,q_2} \},
\]

\[
\Gamma_{ijl_1}^{q_1} = \bigcup_{q_2=1}^{k(i,i,q_1)} \Gamma_{ijp}^{q_1,q_2},
\]

where \( k(i,i,q_1) \) are finite natural numbers; \( \Gamma_{ijl_1}^{q_1} \) and \( \Gamma_{ijl_2}^{q_1,q_2} \) are enclosed totally bounded domains in \( E^{n-1} \); \( \Omega_{ij}^{q_1,q_2} \) is the domain of the function \( A_{ijp}^{q_1,q_2} \) definition; \( A_{ijp}^{q_1,q_2} \) are functions with the same constraints as \( A_{ij}^{q_1}, p = 0 \), if the points of the domain \( \Omega_{ij}^{q_1} \) correspond to the lower values of \( x_{l_2} \) relative to \( \Gamma_{ijp}^{q_1,q_2} \), otherwise, \( p = 1 \).

Having performed the operation of getting the projections \( n \) times, the following ratios are obtained:

\[
x_{l_n} = A_{ijp}^{q_1,q_2, \ldots, q_n}, \quad (7)
\]

delineating domains \( \Omega_{ij}^{q_1,q_2, \ldots, q_{n-1}} \), where

\[
q_s = \frac{k(i,j,q_1, \ldots, q_{s-1})}{s = \overline{2,n}},
\]

\( k(i,j,q_1, \ldots, q_{s-1}) \) are finite natural numbers. It is evident that (7) sets boundary points for segments on the axis \( x_{l_n} \).

Let us denote \( W_s = W_{l_s} W_{l_{s-1}} \ldots W_{l_1} \). Through the fixed point \( x^0 \) let us define the sets:

\[
S_s = \{ x \in E_n | W_s x = W_s x^0 \}, \quad s = \overline{2,n-1},
\]

\[
S_s^p(i, j, q_1, \ldots, q_{s-1}) = \bigcup_{q_s} \Gamma_{i,j,p}^{q_1,q_2, \ldots, q_s}.
\]

**Property 6.** The sets

\[
S_1 \cap S_s^p(i, j, q_1, \ldots, q_{s-1}) \cap S_{s+1}^p(i, j, q_1, \ldots, q_{s-1}),
\]
s = \frac{2}{n} - 1, should have no more than one point for any \( x^0 \).

If, at some stage of obtaining projections, the boundary

\[ \Gamma_{i,j}^{q_1,q_2,...,q_s} \]

satisfies the ratio

\[ (x,y) = 0, \ x \in \Gamma_{i,j}^{q_1,q_2,...,q_s}, \ y \in W_n \mathbb{E}^n, \]

then \( \Gamma_{i,j}^{q_1,q_2,...,q_s}, \ s = \frac{2}{n} - 1 \), there is no need to store, since (8) are restored by minimum and maximum parameter values.

The described representation has the structure of the graph shown in the figure 2. The graph vertex corresponds to the source surface. The second level vertices correspond to the boundaries of the source surface projection, the third level vertices correspond to the boundaries of the projections of the second level structures, etc. At the level \( n \), there are boundary points for intervals on the axis \( x_l \). Each column has a triple \( (i,j,q) \). The full set of such graphs gives the representation of the object as a whole. The graph arcs have the feature \( p \) equal to the value of the feature \( p \) in (6) and (7). Several graphs may have common vertices at the level \( s = \frac{2}{n} \), but the features of arcs coming into them may be different. Based on the definitions introduced, some model properties will be obtained and then used.

![Figure 2. Data view structure](image)

**Theorem 1.** There is one and only one domain \( \Omega_{i,0} \) not bounded in \( E^n \) at \( n \geq 2 \).

Existence. Let us assume that all domains \( \Omega_i, \ i = 1,N \), are bounded, then \( \Omega = \bigcup_{i=1}^{N} \Omega_i \) will be bounded as \( N \) is finite. From the property 1 it follows that \( \Omega = E^n \), but \( E^n \) is not bounded, which contradicts the assumption made.

Uniqueness. Let there be \( N_0 > 1 \) of the domains \( \Omega_j, \ j = 1,N_0 \), not bounded in \( E^n \). Let us consider the sequence of the sets:

\[ S_m(R,a) = \{ x \in E^n | m \times R \leq \| x - a \| \leq (m + 1) \times R \}, \]
where \( R \) is a positive real finite number; \( a \) is a vector in \( E^n \).

\( R \) and \( a \) are selected so that all the bounded domains \( \Omega_i \in S_0(R, a) \). Assuming that all domains are not bounded, \( R = 1 \) and \( a \) is a zero vector in \( E^n \). By definition \( S_m(R, a) \) for any finite \( m \) is closed, bounded and connected. Then \( \Lambda_m(R, a) = \bigcup_{i=0}^{m} S_m(R, a) \) is a closed ball with the radius \((m + 1) \times R \) and the center \( a \), \( E^n = \bigcup_{i=0}^{\infty} S_m(R, a) \). \( S_{m-1}(R, a) \cap S_m(R, a) \) is a closed sphere in \( E^n \) with the radius \( m \times R \) and the center \( a \) for any finite \( m \).

There is such number \( m_1 \) that \( S_{m_1}(R, a) \) contains more than one unbounded domain \( \Omega_j \). If such number cannot be found, it means that the set \( \Lambda_{\infty}(R, a) = E^n \) contains only one unbounded domain, which contradicts the assumption. Thus, \( S_{m_1}(R, a) \) contains several unbounded domains \( \Omega_j \). By the property 2 there is such boundary \( \Gamma_{ij} \) that \( \Gamma_{ij} \cap S_{m_1}(R, a) \neq \emptyset \). Let \( x_1 \in \Gamma_{ij} \cap S_{m_1}(R, a) \). Next such \( m_2 > m_1 \) is found that the set \( S_{m_1}(R, a) \) belongs to more than one unbounded domain \( \Omega_j \). If the numbers \( m_2 \) cannot be determined, then, similarly to the previous reasoning, it follows that there is only one unbounded domain. So there is such \( \Gamma_{ij} \) that \( \Gamma_{ij} \cap S_{m_2}(R, a) \neq \emptyset \). Let us select \( x_2 \in \Gamma_{ij} \cap S_{m_2}(R, a) \). Continuing this process, one will get the sequence \( x_1, x_2, \ldots, x_k, \ldots \), which by the structure has no limit point and is unlimited.

The number of different index pairs \((i, j)\) for boundaries \( \Gamma_{ij} \) is finite. In this case there is the pair \((i_0, j_0)\) for which \( \Gamma_{i_0, j_0} \) contains an infinite number of sequence elements. Since any infinite subsequence \( x_{i_1}, x_{i_2}, x_{i_k}, \ldots \) is unlimited, then \( \Gamma_{i_0, j_0} \) will also be unlimited, which contradicts the property 2. The theorem has been proven.

Note. At \( n = 1 \) it is easy to show that there are two unbounded domains: \( x \to +\infty, x \to -\infty \).

Let us assume that this is one domain with the number \( i_0 \).

**Theorem 2.** The areas of \( \Omega_i, i = 1, N \), except one \( \Omega_{i_0} \), are compact, and the domain \( \Omega_{i_0} \) is locally compact.

**Proof.** The domain \( \Omega_i, i \neq i_0 \) is bounded, therefore any sequence in it will be limited. It should be shown that all the cluster points of the sequence belong to the domain \( \Omega_i \). Let us construct the sequence \( x_{i_1}, x_{i_2}, \ldots, x_{i_k}, \ldots \in \Omega_i \). Due to the boundedness, for any cluster point \( x^* \) one can select fundamental Cauchy subsequence \( x_{i_1}, x_{i_2}, \ldots, x_{i_k}, \ldots \) converging to \( x^* \). Suppose that \( x^* \notin \Omega_i \). Then due to the property 1 there is such \( j \) that \( x^* \in \Omega_j, j \neq i \). For any neighbourhood of the point \( x^* \) there is at least one point belonging to \( \Omega_i \) and at least one point belonging to \( \Omega_j \). By the property 2 \( x^* \in \Gamma_{ij} \). But \( \Gamma_{ij} \subset \Omega_i \), therefore, \( x^* \in \Omega_i \). Consequently, \( \Omega_i \) is closed and compact.

For an arbitrary point \( x \in \Omega_{i_0} \) some of its closed neighbourhood \( S(x) \) will be considered. In a similar way you can show that the set \( S(x) \cap \Omega_{i_0} \) is closed, which proves the local compactness of the domain \( \Omega_{i_0} \). The theorem has been proven.

Consider for \( \Gamma_{ij}^{\alpha} \) a bounding rectangle \( P_{ij}^{\alpha} \), defined by the segments \([a_1^\alpha, a_1^\beta], a_1^\alpha < a_1^\beta \), for \( i \)-th component \( x_i \) of the vector \( x = (x_1, x_2, \ldots, x_n) \in E^n \). Moreover, if \( x \in \Gamma_{ij}^{\alpha} \), then \( x \in P_{ij}^{\alpha} \). Generally speaking, there is a whole family of \( S(P_{ij}^{\alpha}) \) containing rectangles for \( \Gamma_{ij}^{\alpha} \). However, we will consider the minimal one \( P_{ij}^{\alpha} \) for which the differences \( a_i^\alpha - a_i^\beta, i = 1, \ldots, n \), are minimal.

Due to the closedness and boundedness of \( \Gamma_{ij}^{\alpha} \) all \( a_i^\alpha \) and \( a_i^\beta \) exist and are finite.

Consider also the axis \( L \) in \( E^n \) defined by the point \( x^0 \in E^n \) and the finite increment vector \( \beta = (\beta_1, \beta_2, \ldots, \beta_n) \) for each coordinate:

\[
L = \{x \in E_n | x = x^0 + \alpha \beta, \alpha \in R, \alpha \geq 0\}. \tag{9}
\]

Moreover, there is \( i \), for which \( \beta_i \neq 0 \).

The axis \( L \) may have a non-empty intersection with \( P_{ij}^{\alpha} \), \( L(i, j, q_1) = L \cap P_{ij}^{\alpha} \). It is clear that \( L(i, j, q_1) \) is closed and bounded, if \( L(i, j, q_1) \neq \emptyset \). Let us consider the set \( M^k(i, j, q_1) \subset L(i, j, q_1) \),
consisting from \( k \) points \( x^{(1)}, x^{(2)}, \ldots, x^{(k)} \), where \( x^{(1)} \) and \( x^{(k)} \) are the ends of the segment \( L(i,j,q_1) \). The set \( M^{k}(i,j,q_1) \) forms partitioning of \( L(i,j,q_1) \) into \( k-1 \) intervals with the same length: \( \|x^{(i)} - x^{(i-1)}\| = \|x^{(i+1)} - x^{(i)}\|, i = 2, k-1 \), where \( k > 2 \) is a finite integer.

**Definition 1.** The domain \( \Omega_i \) will be called \( k \)-fuzzy, if the following is satisfied:

\[
L(i,j,q_1) \neq \emptyset, \quad L \cap (\Omega_i \setminus \bigcup_{(i,j)} \Delta_{ij}) \neq \emptyset, \quad (10)
\]

then

\[
\Omega_i \cap M^{k}(i,j,q_1) \neq \emptyset, \quad (11)
\]

where \( i \) and \( q_1 \) take all the values from the existing triples \((i,j,q_1)\) and \((j,i,q_1)\) with the fixed number \( i \), of the domain \( \Omega_i \).

Let us get the estimate of the value \( k \).

The rectangle \( P_{ij}^{q_1} \) in \( E^n \) is defined by the pair of vectors: \( \alpha^1 = (\alpha^1_1, \alpha^1_2, \ldots, \alpha^1_n), \alpha^2 = (\alpha^2_1, \alpha^2_2, \ldots, \alpha^2_n) \). It is clear that maximum length of the segment \( L(i,j,q_1) \) equals \( \|\alpha^2 - \alpha^1\| \). Let us denote by \( \Gamma\Delta_{ij} \) the boundary of the domain \( \Delta_{ij} \) and \( r_{ij} = \min||\Gamma_{ij} - \Gamma\Delta_{ij}|| \).

**Theorem 3.**

\[ k \geq \frac{\|\alpha^2 - \alpha^1\|}{r_{ij}} + 1. \]

Proof. Let us assume that the conditions of the definition are not fulfilled for the given value \( k: L(i,j,q_1) \neq \emptyset, S = L \cap (\Omega_i \setminus \bigcup_{(i,j)} \Delta_{ij}) \neq \emptyset, \) however \( \Omega_i \cap M^{k}(i,j,q_1) = \emptyset \). Choose an arbitrary point \( x \in S \). Then there are such points \( x^{(m)} \) and \( x^{(m+1)} \in M^{k}(i,j,q_1) \) that \( x^{(m)} \notin \Omega_i \) and \( x^{(m+1)} \notin \Omega_i \), but \( x \in [x^{(m)}, x^{(m+1)}] = L_m \), where \( L_m \) is a sector on the axis \( L \). The length of the segment \( L_m \) equals \( r_m = \|x^{(m+1)} - x^{(m)}\| \leq \frac{\|\alpha^2 - \alpha^1\|}{k-1} \). We will get the lower boundary \( r_m \). Due to the definition, \( S \) is open, so there is a neighbourhood of the point \( x \in S \) with the radius \( R > 0 \) also belonging to \( S \). Since the ends of \( L_m \) are outside of \( \Omega_i \) and \( L_m \) has a non-empty intersection with \( S \), then according to the property 2 there exists the point \( x^* \in \Gamma_{ij} \cap L \). The point \( x^* \) has a closed neighbourhood \( L_{m} \) with the radius \( r_m' \geq r_{ij} \) by the definition \( r_{ij} \).

It is evident that \( L_{m}' \) has an empty intersection with the considered neighbourhood of the point \( x \). Then \( r_m \geq r_m' + 2R \) and

\[
k - 1 \leq \frac{\|\alpha^2 - \alpha^1\|}{r_m} \leq \frac{\|\alpha^2 - \alpha^1\|}{r_{ij}} < \frac{\|\alpha^2 - \alpha^1\|}{r_m' + 2R} \leq k - 1.
\]

It follows that \( k < k \). The resulting contradiction proves the theorem.

Consider another important property of \( k \)-fuzzy domains.

**Theorem 4.** If the domain \( \Omega_i \) is \( k \)-fuzzy, then it is \((k+1)\)-fuzzy.

Proof. Let us suppose the opposite. Then there are \( \Gamma_{ij}^{q_1} \) and the axis \( L \) defined by the point \( x^0 \) and the vector of relative increments \( \beta \), such that

\[
S = L(i,j,q_1) \cap (\Omega_i \setminus \bigcup_{(i,j)} \Delta_{ij}) \neq \emptyset,
\]

\[
M^{k+1}(i,j,q_1) \cap \Omega_i = \emptyset.
\]
Let us randomly choose the point \( x^* \in S \). The segment \( L(i, j, q_1) \) is broken down by the points \( x^{(s)} \in M^{k+1}(i, j, q_1), s = 1, k + 1 \), into \( k \) equal segments. Two situations are possible:

1. \( x^* \in \bigcup_{s=1}^{k-1} [x^{(s)}, x^{(s+1)}] \).
2. \( x^* \in [x^k, x^{k+1}] \).

In the first case, the axis \( \mathcal{L} \) is constructed, for which \( \bar{x}^0 = x^{(k)} \) and \( \bar{\beta} = -\beta \). Clearly \( \mathcal{M}^k(i, j, q_1) \) for \( \mathcal{L} \) coincides with the set \( M^{k+1}(i, j, q_1) \setminus x^{(k+1)} \). Based on the assumption, it follows that \( \mathcal{M}^k(i, j, q_1) \cap \Omega_i = \emptyset \) while \( \mathcal{L} = \mathcal{L}(i, j, q_1) \cap (\Omega_i \setminus \bigcup_{(i,j)} \Delta_{ij}) \neq \emptyset \) since \( S \) at least contains \( x^* \). This indicates that \( \Omega_i \) is not \( k \)-fuzzy.

In the second case, \( \mathcal{L} \) is set by the point \( \bar{x}^0 = x^{(2)} \) and the vector \( \bar{\beta} = \beta \). Then, \( \mathcal{M}^k(i, j, q_1) \) for \( \mathcal{L} \) coincides with the set \( M^{k+1}(i, j, q_1) \setminus x^{(1)} \) for \( L \). As in the first case, it means that \( \Omega_i \) is not \( k \)-fuzzy.

It should be noted that both cases cannot be performed simultaneously, as it involves: \( x^* = x^{(k)} \), contradicting the assumption \( M^{k+1}(i, j, q_1) \cap \Omega_i = \emptyset \). The theorem has been proven.

Note. The property proved in the previous theorem allows one to choose the universal value \( k \) for all domains \( \Omega_i, i = 1, N \), which will make \( k \) independent from triples \((i, j, q_1)\).

Conclusion
The proposed data model is universal in terms of describing domains in multidimensional Euclidean space. Its application field is not limited by data analysis. These can be CAD systems in civil engineering, systems in geological exploration, etc. The question of calculating domain boundaries is out of the scope of this paper. The most suitable in this case are methods of mixed surface approximation [10]. However, classical methods of approximation should be adapted to this task (it is necessary to modify the optimization criterion), because in space, not the points of surfaces are set, but the points from which the surfaces should be distanced significantly.

Acknowledgments
The work was supported by the program of fundamental scientific researches of the SB RAS No I.5.1., project No 0314-2019-0020.

References
[1] Xi D and Gallo P 2019 An additive boundary for group sequential designs with connection to conditional error Statistics in Medicine 38 23 pp 4656–69
[2] Oliveira F L P, Cançado A L F, de Souza G, Moreira G J P and Kulldorff M 2018 Border analysis for spatial clusters International Journal of Health Geographics 17 5
[3] Kong Q L and Zhu Q M 2007 Incremental procedures for partitioning highly intermixed multi-class datasets into hyper-spherical and hyper-ellipsoidal clusters Data \& Knowledge Engineering 63 2 pp 457–477
[4] Kashirina I L, Fedutinov K A, Azarnova T V and Bondarenko Iu V 2019 System analysis of monitoring ecological information using Fuzzy ART neural network Journal of Physics: Conference Series 1203 012090
[5] Newman W M and Sproull R F 1973 Principles of interactive computer graphics (New York : McGraw-Hill) p 570
[6] Kaiser A, Zepeda J A Y and Boubekeur T 2019 A survey of simple geometric primitives detection methods for captured 3D data Computer Graphics Forum 38 1 pp 167–196
[7] Li D, Yang D Y and Liu G B 2019 Research on Key Technologies of Night Detection of License Plate Recognition System 2nd International Conference on Mechanical, Electronic and Engineering Technology (MEET-2019) pp 101–105
[8] Geist M and Pietquin O 2011 Parametric value function approximation: A unified view IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL) 11–15 April 2011 pp 9–16
[9] Fusch P, Fusch G E and Ness L R 2018 Denzin’s Paradigm Shift: Revisiting Triangulation in Qualitative Research Journal of Social Change 10 1 pp 19–32
[10] Zykin S V 1986 Mixed approximation *Mathematical models for the representation of information and data processing problems* (Akad. Nauk SSSR Sibirsk. Otdel., Vychisl. Tsentr, Novosibirsk) pp 72–79 MR952417