ISSN 0005-1179, Automation and Remote Control, 2019, Vol. 80, No. 1, pp. 112–123. © Pleiades Publishing, Ltd., 2019.
Russian Text © A.L. Myachin, 2019, published in Avtomatika i Telemekhanika, 2019, No. 1, pp. 138–152.

OPTIMIZATION, SYSTEM ANALYSIS, AND OPERATIONS RESEARCH

Pattern Analysis in Parallel Coordinates Based on Pairwise Comparison of Parameters

A. L. Myachin[*,**]

[*] *National Research University Higher School of Economics, Moscow, Russia*
[**] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: amyachin@hse.ru

Received October 17, 2017
Revised April 2, 2018
Accepted November 8, 2018

Abstract—We present the basic properties of the a new pattern analysis method in parallel coordinates; results of the method do not depend on the ordering of data in the original sample of objects being analyzed. We prove that clusters obtained with this method do not overlap. We also show the possibility of representing objects of one cluster in the form of monotonically increasing/decreasing functions.

Keywords: pattern analysis, ordinal-invariant pattern clustering, cluster analysis

DOI: 10.1134/S0005117919010107

1. INTRODUCTION

With the development of information technology and accumulation of large amounts of data, it becomes increasingly important to develop methods that automate the process of recognizing and isolating various groups with similar properties. The first such works include [1], which suggests to split a set of three types of irises according to four features using a linear discriminant algorithm. The database of 150 irises described in [1] is still one of the most popular in machine learning, and the methodology proposed by R. Fisher is discussed in detail in a number of textbooks (see, e.g., [2, 3]). This direction was further developed in cluster analysis algorithms based on applying different measures of proximity between objects [4, 5]. However, the need to study increasingly complex processes (economic, financial, social, and other) has led to the need to take into account not only the proximity of values but also the proximity of the data structures themselves, which, in turn, has led to the development of methods for their visual presentation and analysis (the field this work, broadly speaking, belongs to).

Of the more recent studies that take into account these trends, we highlight, in particular, applied research in the banking sector [6, 7] (analysis of CAMEL indicators [8–10]), management [11], and macroeconomics [12]. This direction of study is based on the concept of a "pattern," which is defined differently in different areas of knowledge. Basic definitions have been given in [13]:

1) "as the essence of a phenomenon that has repeating features";

2) "as a property of repeating components united by a common structure";

3) "as a process that fixes the model of interaction of the objects in question, including repetitions."

In [2], a pattern is understood as "any relation, dependency or structure inherent in some dataset," and analysis of patterns is defined as "the process of finding common relationships in a dataset." Based on these definitions, by a pattern in this work we will mean a combination of certain qualitatively similar features. It has been noted in [2] that in order to recognize a pattern analysis method as effective, its implementation must satisfy the following conditions: ability to process large amounts of data (which implies relatively low computational complexity), robustness, and statistical stability. In this work, we take into account one more criterion mentioned in [14]: the independence of the final results on the choice of the initial sequence of parameters. In this regard, we consider pattern analysis methods proposed in [15]: ordinal-fixed and ordinal-invariant pattern clustering. In this work we generalize the algorithms that implement them, propose methods for pattern recognition and combining similar objects into groups, and explore a number of properties of these groups, including convenient methods for visualizing them and defining "average/central" objects in a cluster. Thus, the goal of the work is to summarize and introduce structure to the algorithms developed for pattern analysis whose final result is independent of the input data ordering, and also to study the basic properties of both the algorithms and the groups of objects obtained from them. For convenience and integrity of the presentation, proofs of all propositions formulated in this work are relegated to the Appendix.

2. PATTERN ANALYSIS: BASIC CONCEPTS

We investigate a set consisting of m objects characterized by n parameters. We denote elements of the set by x_i; a set of parameters of a specific object, by $x_i = (x_{i1}, \ldots, x_{ij}, \ldots, x_{in})$, where x_{ij} denotes the jth indicator of the ith object. The main problem is to find and unite qualitatively similar objects. For visualization, we use parallel coordinates system [14, 16, 17], which usually consists of vertical lines (axes) reflecting values of the parameters. On these lines we mark their actual values, and then the values are connected by segments. As a result, we get polylines that characterize the analyzed objects.

Let us explain the idea of the method of combining objects into separate groups (clusters) by the form of polylines with a hypothetical example.

Example. We consider five objects characterized by four parameters (A, B, C, and D), whose values are listed in Table 1.

Objects 1–5 are visualized on Fig. 1.

The figure clearly shows that objects 1 and 2 have similar structures. This is also true for objects 4 and 5. We consider object 3 separately. Despite the fact that absolute values of the

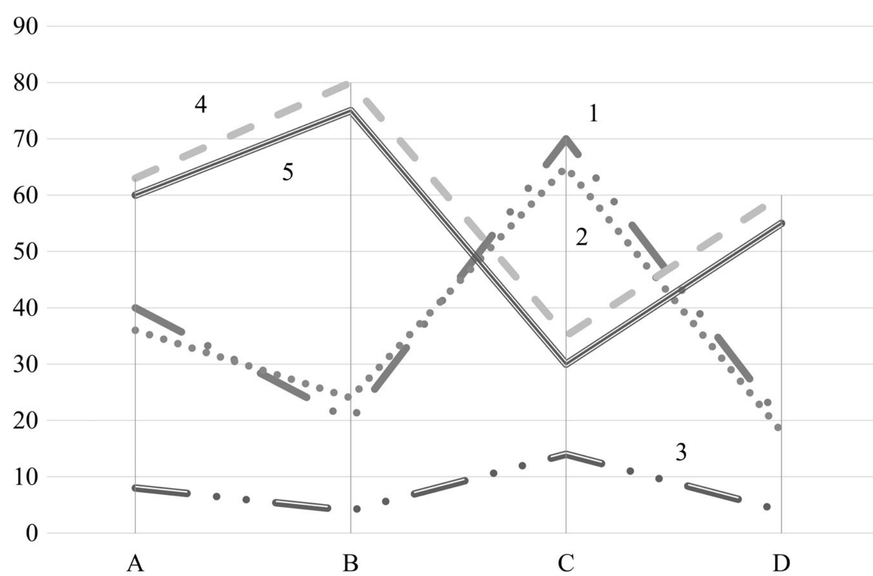


Fig. 1. Sample hypothetical objects.

Table 1. Example with hypothetical objects

Objects	A	B	C	D
1	40	20	70	20
2	36	24	65	18
3	8	4	14	4
4	63	80	35	60
5	60	75	30	55

parameters of this object are very different from objects 1 and 2, their structures are similar: parameters of object 1 are parameters of object 3 multiplied by 5. Thus, all three objects are described by polylines of the same structure (the same pattern), which gives a justification to uniting them by this criterion.

This approach has proved useful for a wide range of applications. In particular, the work [6] presented a dynamic analysis of the patterns of parameters of 1018 Russian banks for 1999–2003. The original data is based on the fundamental characteristics of banks, CAMEL parameters (C—capital sufficiency, A—asset quality, M—management, E—profit, L—liquidity), with some additions.

On the basis of quarterly data for the specified period, 19 342 polylines were constructed; they allowed to identify 151 typical patterns, with the first 50 covering 90.14% of all data, and the first 13 covering 52.19%. In [7], a similar approach was applied to a study of 55 Turkish banks. The raw data was also based on CAMEL, and as a result 27 patterns were formulated.

The pattern analysis method in parallel coordinates system has proven its effectiveness for different applications [13, 16]; however, the method itself is sensitive to the sequence of parameters. Below we present algorithms for its implementation based on pairwise comparisons of the parameters of the ordinal-fixed and ordinal-invariant pattern clustering, which makes it possible to eliminate this drawback.

3. ORDINAL-FIXED AND ORDINAL-INVARIANT PATTERN CLUSTERING

Let us give a brief description of the ordinal-fixed and ordinal-invariant pattern clustering algorithms that unite objects into groups/clusters by the form of their in parallel coordinates (a full description of the methods is given in [15, 18]).

As noted above, the input data is represented by a set X consisting of m objects, each of which is characterized by n parameters. We construct the coding c_i of each object $x_i \in X$ by pairwise comparison of its parameters according to the formula

$$c_i = \sum_{s=1}^{n} 10^{s-1} z_i^{n-s}, \tag{1}$$

where z_i^s is defined as

$$z_i^s = \begin{cases} 1, & \text{if } x_{is} < x_{is+1} \\ 0, & \text{if } x_{is} = x_{is+1} \\ 2, & \text{if } x_{is} > x_{is+1}. \end{cases} \tag{2}$$

If $c_i = c_k$, objects x_i and x_k are combined into a single cluster, if not, are separated. Note that the value of z_i^s is defined in expression 2 by comparing adjacent values of x_{is} and x_{is+1} in a given sequence. Therefore, in what follows we call this method ordinal-fixed pattern clustering, and the clusters found with this method will be called ordinal-fixed pattern clusters.

The computational complexity N_{fix} of this method can be computed as

$$N_{fix} = \frac{m^2(m-1)(n-1)}{2}.$$

Since N_{fix} is relatively low, this method is convenient to use for preliminary data analysis.

The second approach, ordinal-invariant pattern clustering, was developed in order to make the results independent of the original sequence of parameters. This method is needed, as noted above, for the reasons remarked upon in some works on parallel coordinates (including [14]); these remarks indicate the need for "extremely careful selection of the sequence of analyzed parameters," because it affects the appearance of new patterns and, in the general case, a different sequence can lead to a different character of polylines and, as a result, to different results.

The ordinal-invariant pattern clustering algorithm is similar to ordinal-fixed and is based on a comparison of object encodings. The objects themselves in this case are presented as complete directed weighted graphs, whose vertices correspond to the parameters in questions, and values of the edges connecting them are the results of pairwise comparisons (we assume comparisons of the form "greater than," "equal," and "less than"). Using the values of the edges, we construct an additional object encoding:

$$c_i^{dop} = \sum_{s=1}^{n-2} \sum_{j=s+2}^{n} 10^{j-(s+2)} e_i^{sj}, \tag{3}$$

where e_i^{sj} is the value of the edge of the graph connecting vertices s and j, which, similar to (2), is defined by the formula

$$e_i^{sj} = \begin{cases} 1, & \text{if } x_{is} < x_{is+1} \\ 0, & \text{if } x_{is} = x_{is+1} \\ 2, & \text{if } x_{is} > x_{is+1}. \end{cases} \tag{4}$$

The principle of grouping is similar to ordinal-fixed pattern clustering: if $c_i^{dop} = c_k^{dop}$ then objects x_i and x_k are combined into a single cluster, otherwise they are divided into different clusters.

In what follows, we call a cluster constructed through ordinal-invariant pattern clustering an ordinal-invariant pattern cluster.

The computational complexity N_{inv} of this algorithm is defined as

$$N_{inv} = \frac{m^2 n(m-1)(n-1)}{4};$$

however, using Statement 2 below, it can be reduced to the complexity of the sorting algorithm.

Next, we find the maximum number of ordinal-fixed and ordinal-invariant pattern clusters.

In the considered methods, the criterion for placing objects into a single cluster is the fact that codes formed in the pairwise comparisons of the corresponding object parameters coincide. Using a well-known expression from coding theory (with a coding alphabet consisting of β different characters and the length of the code sequence γ, the maximum number V of different code combinations is β^γ) , we get

$$V = R^N,$$

where R is the number of possible values characterizing the results of pairwise comparisons of the corresponding object parameters (defined by formulas (2) and (4)).

Since for ordinal-fixed pattern clustering the code is formed by the $n-1$ pairwise comparisons, the number of possible ordinal-fixed pattern clusters V_{fix} can be found as

$$V_{fix} = R^{(n-1)} = 3^{(n-1)},$$

and the number of clusters formed as a result of ordinal-invariant pattern clustering V_{inv} is

$$V_{inv} = R^{\frac{n(n-1)}{2}} = 3^{\frac{n(n-1)}{2}}.$$

Taking the ratio of these values, we obtain an estimate of the ratio between the maximum possible numbers of ordinal-invariant and ordinal-fixed pattern clusters:

$$\frac{V_{inv}}{V_{fix}} = \frac{R^{\frac{n(n-1)}{2}}}{R^{(n-1)}} = 3^{\frac{n(n-1)}{2}}.$$

Remark 1. The resulting expression determines the maximum possible number of subclusters that can be distinguished within a ordinal-fixed pattern cluster. Their actual number may be significantly smaller. In particular, if a ordinal-fixed pattern cluster satisfies the conditions of Statement 2 (see Section 4), then it is an ordinal-invariant pattern cluster itself, and by virtue of Statement 1 (see below) it does not contain any other subclusters.

4. ORDINAL-INVARIANT PATTERN CLUSTERING: BASIC PROPERTIES

We present three statements that demonstrate important properties of the method described above.

Statement 1. *Clusters obtained with ordinal-invariant pattern clustering do not overlap.*

The proof is given in the Appendix.

The following remark can be made.

Remark 2. Statement 1 is very important because it affirms the fact that results of order-invariant pattern clustering are unambiguous. This means that an arbitrary arrangement of objects in the original set, as well as an arbitrary order of their parameters, and the use of ordinal-invariant pattern clustering for an arbitrary number of times on the same data does not affect the result of clustering.

Note that if the order of the parameters is strictly fixed and does not change, then Statement 1 also holds for ordinal-fixed pattern clusters.

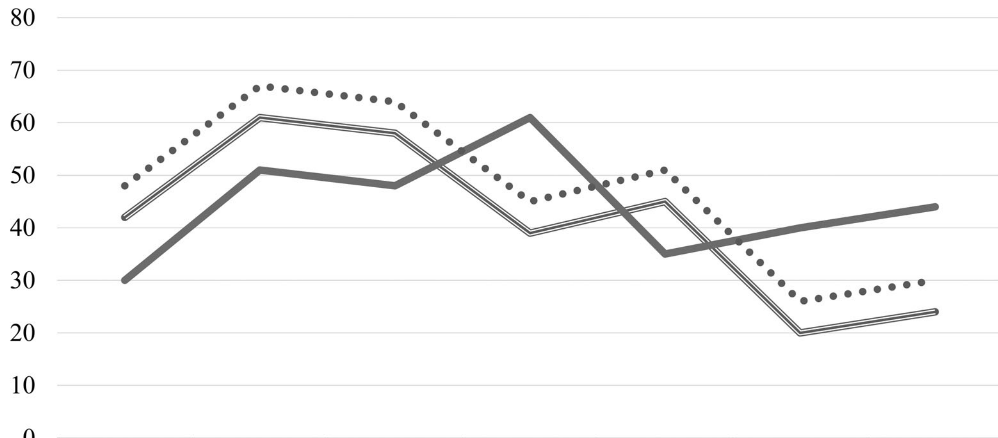
Statement 2. *If there exists a sequence of parameters for which their values form a strictly monotonically increasing/decreasing sequence for every object of the original set X, then this set is an ordinal-invariant pattern cluster.*

The proof is given in the Appendix.

Remark 3 is in order here.

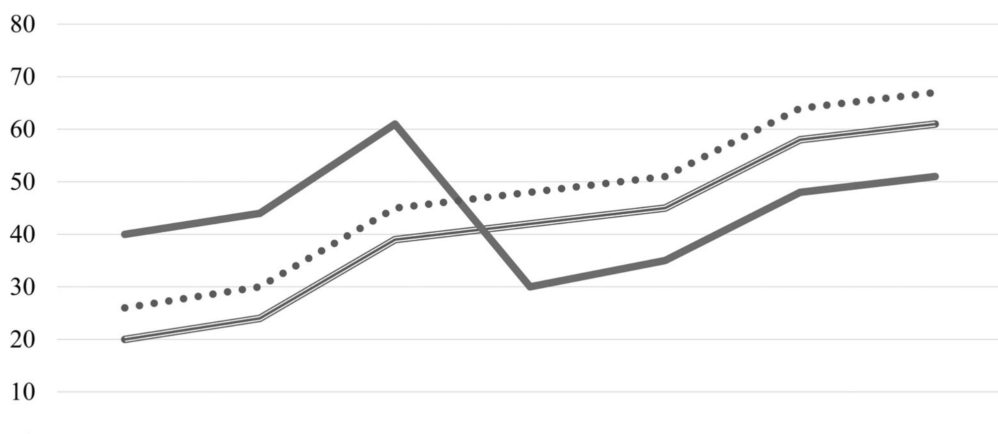
Remark 3. It is convenient to use the property of ordinal-invariant pattern clusters determined by Statement 2 for preliminary comparison of different groups of objects and visual perception of their distinctive features. As an example, we consider a set of three objects whose parameters and the corresponding polylines are shown in Fig. 2. Using the methods described above allows us to distinguish two ordinal-invariant pattern clusters: {Object 1; Object 3} and {Object 2}.

We arrange the parameters in such a way that their values form a monotonically increasing sequence for the first and third objects that comprise an ordinal-invariant pattern cluster (see Fig. 3). In this case, it is easier to see in what way they differ from object 2.



	A	B	C	D	E	F	G
1	42	61	58	39	45	20	24
2	30	51	48	61	35	40	44
3	48	67	64	45	51	26	30

Fig. 2. A hypothetical example of two ordinal-invariant pattern clusters.



	F	G	D	A	E	C	B
1	20	24	39	42	45	58	61
2	40	44	61	30	35	48	51
3	26	30	45	48	51	64	67

Fig. 3. A hypothetical example of two ordinal-invariant pattern clusters.

Statement 3. *For objects of an ordinal-invariant pattern cluster, there exists an order of parameters for which their values form a monotonous non-decreasing/non-increasing sequence for every object in the cluster.*

The proof is given in the Appendix.

We conclude this section with Remark 4.

Remark 4. An important consequence of Statements 2 and 3 is the possibility of using them to reduce the computational complexity of ordinal-invariant pattern clustering to the complexity of the sorting algorithm.

5. USING ORDINAL-INVARIANT PATTERN CLUSTERING IN THE STUDY OF ECONOMIC, INNOVATIVE, AND EDUCATIONAL PARAMETERS IN THE RUSSIAN FEDERATION

To illustrate the proposed methods, as well as some basic properties, we consider an example of real data from [19], which studies the parameters of scientific, educational, and innovation activity of regions of the Russian Federation for 2007–2010. On the basis of the Russian regional innovation index [20], 6 blocks of parameters were constructed: socio-economic conditions (A), educational potential (B), potential for research and development activity (X), the impact of research and development (C), potential for innovation (D), and performance of innovation activities (E). As a result of correlation analysis, the scientific-technical potential block was excluded, and regions of the Russian Federation were subdivided on the basis of blocks A–E.

To begin with, we use ordinal-invariant pattern clustering for partitioning regions into clusters. The result is 22 clusters containing more than 5 objects, 12 clusters with 3 to 5 objects, and 1 cluster with a "unique object." Sample clusters are shown in Fig. 4.

Next, we take one of the clusters obtained in [20] and check whether it can be attributed to ordinal-invariant pattern clusters (the data are given in Table 2, with values rounded to the second decimal place).

Table 2. Example with hypothetical objects

Region/year	A	B	C	D	E
Moscow 2009	0.36	0.47	0.91	0.3	0.07
Moscow 2010	0.35	0.42	0.89	0.21	0.11
St. Petersburg 2007	0.34	0.49	0.68	0.3	0.07
St. Petersburg 2009	0.45	0.51	0.7	0.21	0.19
Primorskii Krai 2008	0.27	0.35	0.6	0.24	0.21
Primorskii Krai 2009	0.35	0.36	0.67	0.25	0.21

Each region in Table 2 corresponds to a unified encoding obtained with formulas (1) and (3): "1122" and "222221" (additional encoding). Therefore, the pattern obtained in [19] is an order-invariant pattern cluster. For these regions, we construct polylines, as shown on Fig. 5.

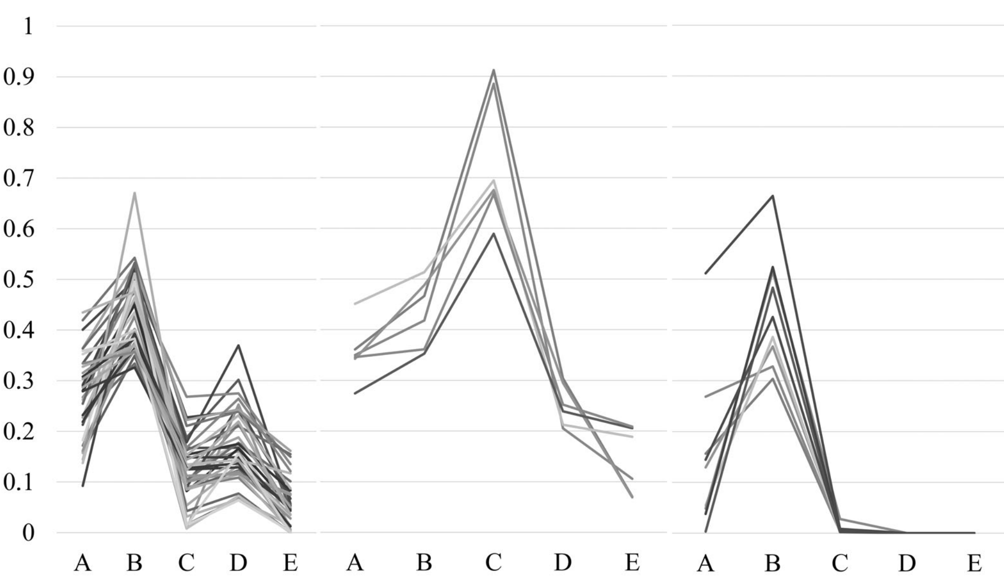


Fig. 4. Sample clusters obtained using ordinal-invariant pattern clustering.

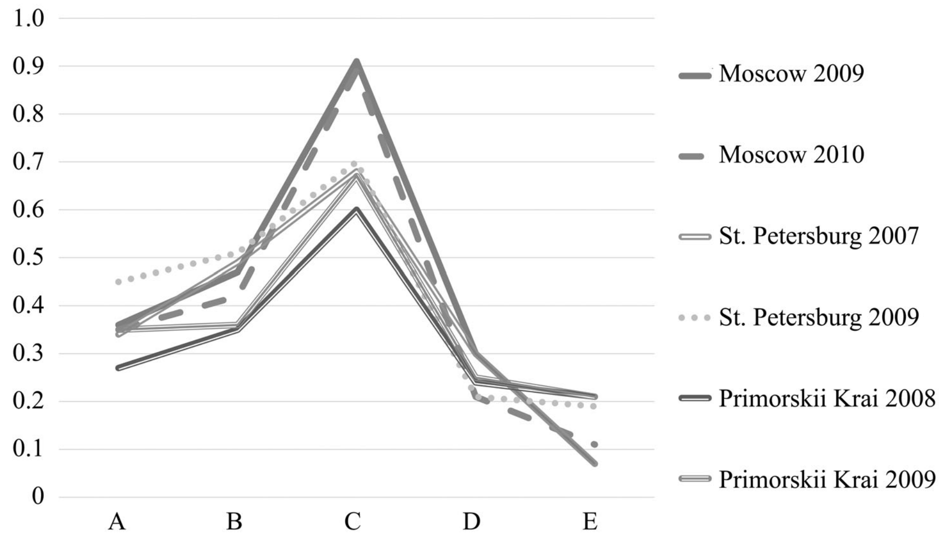


Fig. 5. Polylines of objects corresponding to the order of parameters A, B, C, D, E.

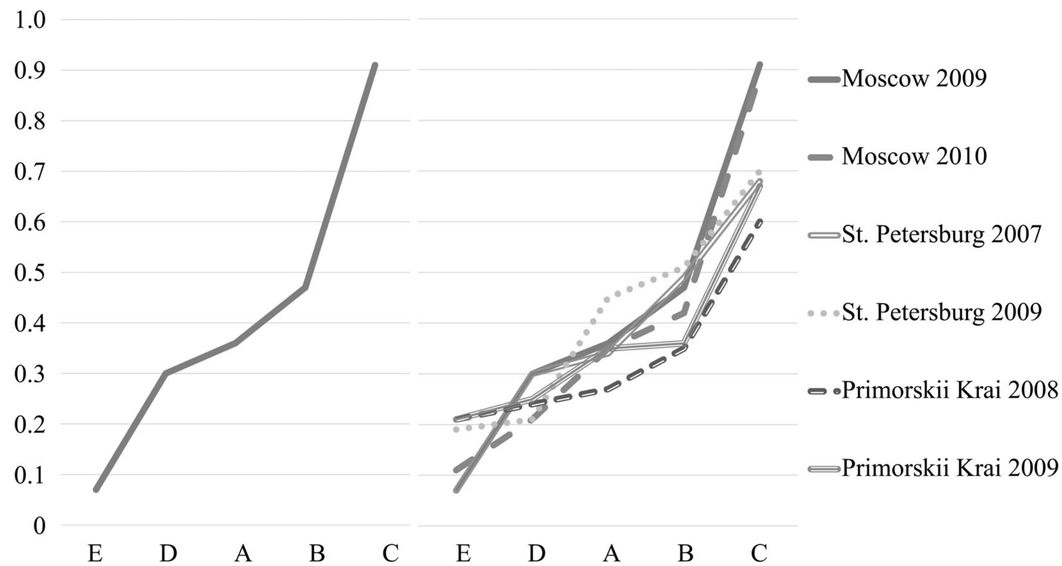


Fig. 6. Polylines of objects corresponding to the order of parameters A, B, C, D, E.

Next we illustrate the validity of Statement 3. We choose the first row of Table 2: (0.36; 0.47; 0.91; 0.3; 0.07). This row describes selected parameters of Moscow in the year 2009. We arrange the values of the parameters in ascending order (E, D, A, B, C). This order of parameters corresponds to a pattern that has the form of non-decreasing polyline. With this arrangement of parameters, according to Statement 3 polylines of other objects will also have the form of a non-decreasing function, as shown in Fig. 6. It is easy to check that in this example Statement 2 also holds: for any of the three regions under consideration (for two years each), there is a sequence of parameter positions (E, D, A, B, C) for which they form a strictly increasing sequence, and, therefore, the regions under consideration form a single ordinal-invariant pattern cluster.

6. OPERATIONS ON OBJECTS IN CLUSTERS AND THEIR PROPERTIES

Let us consider some mathematical operations on the objects of ordinal-invariant pattern clusters.

1. Summation. By the sum of two objects $x_1 = (x_{11},\dots, x_{1j}, \dots, x_{1n})$ and $x_2 = (x_{21}, \dots, x_{2j}, \dots, x_{2n})$ we will understand a new object $x_s = x_1 + x_2$, whose parameter values are defined as the sum of the corresponding parameters of objects x_1 and x_2: $x_s = (x_{11} + x_{21}, \dots, x_{1j} + x_{2j}, \dots, x_{1n} + x_{2n})$.

2. Multiplication of an object by a number. The product of an object $x_1 = (x_{11}, \dots, x_{1j}, \dots, x_{1n})$ of an ordinal-invariant pattern cluster and a real number α is a new object $x_\alpha = \alpha x_1$, whose parameter values are defined as products of the parameters of the original object x_1 and the number α: $x_\alpha = (\alpha x_{11}, \dots, \alpha x_{1j}, \dots, \alpha x_{1n})$.

Statement 4. *If two objects are $x_1 = (x_{11}, \dots, x_{1j}, \dots, x_{1n})$ and $x_2 = (x_{21}, \dots, x_{2j}, \dots, x_{2n})$ belong to the same ordinal-invariant pattern cluster, then their sum $x_s = x_1 + x_2 = (x_{11} + x_{21}, \dots, x_{1j} + x_{2j}, \dots, x_{1n} + x_{2n})$ also belongs to this cluster.*

The proof is given in the Appendix.

Statement 5. *If object $x_1 = (x_{11}, \dots, x_{1j}, \dots, x_{1n})$ belongs to some ordinal-invariant pattern cluster v_a^{inv}, then for any positive value α ($\alpha > 0$) the object $x_\alpha = \alpha x_1$ also belongs to this cluster.*

The proof of this Statement is similar to the proof of Statement 4.

Next we formulate important corollaries of Statements 4 and 5.

Corollary 1. *If objects $x_1, \dots, x_i, \dots, x_n$ belong to the same ordinal-invariant pattern cluster, then the "average" ("central") object x_{st} of the form*

$$x_{st} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

also belongs to this cluster.

Corollary 2. *If objects $x_1, \dots, x_i, \dots, x_n$ belong to the same ordinal-invariant pattern cluster, then their linear combination x_{lc} of the form*

$$x_{lc} = \sum_{i=1}^{n} \lambda_i x_i,$$

where $\lambda_i > 0 \mid i = 1, \dots, n$, also belongs to this cluster.

7. CONCLUSION

In this work, we have described two variations of a promising method for data analysis, namely pattern analysis, and described some of their properties. We have found the computational complexity of the corresponding methods. We have formulated and proved five statements and two corollaries, including the disjointness of ordinal-invariant pattern clusters, representation of objects in these clusters as monotonically increasing/decreasing functions, as well as finding the "average" ("central") object in a cluster.

ACKNOWLEDGMENTS

This paper was prepared as a result of research conducted within the framework of the Basic Research Program of the National Research University Higher School of Economics (HSE) and with the use of subsidy funds of Governmental Support of Leading Universities of the Russian Federation "5-100." The author is grateful to Sc.D., Prof. F.T. Aleskerov for assistance in writing this paper.

APPENDIX

Proof of Statement 1. Suppose that there are two different (non-identical) clusters $v_a^{inv} \neq v_b^{inv}$, obtained as a result of ordinal-invariant pattern clustering of some set of objects X. According to Statement 1, $v_a^{inv} \cap v_b^{inv} = \varnothing$.

We prove this statement by contradiction. Suppose that this statement fails, and $v_a^{inv} \cap v_b^{inv} \neq \varnothing$. This means that there exists at least one object $x_i^* \in X : x_i^* \in v_a^{inv} \cap v_b^{inv}$.

Since $v_a^{inv} \neq v_b^{inv}$, there also exists a certain sequence of initial parameters for which the polylines of these clusters have a different form. We denote this sequence by Y.

According to the definition (construction algorithm) of ordinal-invariant pattern clusters, the form of polylines of all cluster objects v_a^{inv} must match for any sequence of initial parameters (including the sequence Y). Since $x_i^* \in v_a^{inv}$, the form of polylines of all objects in a cluster v_a^{inv} is similar to the form of polylines of object x_i^*. Repeating this reasoning for the cluster v_b^{inv}, we conclude that the form of polylines of all objects in cluster v_b^{inv} is also similar to the form of the polylines of object x_i^*. Thus, we come to the conclusion that the form of polylines of all objects in clusters v_a^{inv} and v_b^{inv} is the same (similar to the form of the polyline of object x_i^*). However, in this case, according to the definition (construction algorithm) of ordinal-invariant pattern clustering, clusters v_a^{inv} and v_b^{inv} should be combined into a single ordinal-invariant pattern cluster, which contradicts the assumption that v_a^{inv} and v_b^{inv} are different, non-identical clusters. Therefore, $v_a^{inv} \cap v_b^{inv} = \varnothing$, which proves Statement 1.

Proof of Statement 2. Consider the case of a monotonically increasing sequence of parameters (a similar proof can be given for a monotonically decreasing sequence). Let us prove Statement 2 by induction, i.e., we first verify its validity for a set of two and three objects and then, assuming it holds for $k > 3$ objects, we prove it for $k + 1$ objects.

1. Suppose that the set X contains only two objects: $x_1 = (x_{11}, \ldots, x_{1j}, \ldots, x_{1n})$ and $x_2 = (x_{21}, \ldots, x_{2j}, \ldots, x_{2n})$, and $x_{11} < \ldots < x_{1j} < \ldots < x_{1n}$ and $x_{21} < \ldots < x_{2j} < \ldots < x_{2n}$. Based on the ordinal-invariant pattern clustering method described above, for pairwise comparison of all parameters of a single object we need $n(n-1)/2$ comparisons, and their result is uniquely determined by expression (2). Since parameter values of the first object are sorted in ascending order ($x_{11} < \ldots < x_{1j} < \ldots < x_{1n}$), we get that $x_{1p} < x_{1q} \;\; \forall\, p < q$ and $x_{1r} > x_{1s} \;\; \forall\, r > s$. For object x_2 we arrive at the same conclusion: $x_{2p} < x_{2q} \;\; \forall\, p < q$ and $x_{2r} > x_{2s} \;\; \forall\, r > s$. Since, as we have noted, inequalities for the first and second object are the same, the comparison results determined by expression (4) are also the same, and, as a result, the code sequences defined by expression (3) are identical as well.

Since we have obtained the same encodings for objects x_1 and x_2, we can conclude that these objects can be combined into a single ordinal-invariant pattern cluster.

2. We supplement the analysis with object $x_3 = (x_{31}, \ldots, x_{3j}, \ldots, x_{3n})$, and $x_{31} < \ldots < x_{3j} < \ldots l < x_{3n}$.

Since condition $x_{31} < \ldots < x_{3j} < \ldots < x_{3n}$ is satisfied for the parameters of object x_3, item 1 of the proof implies that

(a) x_1 and x_3 belong to the same ordinal-invariant pattern cluster;

(b) x_2 and x_3 belong to the same ordinal-invariant pattern cluster.

According to Statement 1, clusters obtained using ordinal-invariant pattern clustering do not intersect. Thus, all three objects form a single ordinal-invariant pattern cluster.

3. Assuming that Statement 2 holds for the case of k objects, we verify its validity for the case of $k + 1$ objects. For this purpose, we add to this set a new object $x_{(k+1)} = (x_{(k+1)1}, \ldots, x_{(k+1)j},$

$\ldots, x_{(k+1)n})$, where $x_{(k+1)1} < \ldots < x_{(k+1)j} < \ldots < x_{(k+1)n}$. Similar to item 1, this object forms a single ordinal-invariant pattern cluster with any of the k previous objects. Now Statement 1 implies that all $k+1$ objects belong to the same ordinal-invariant pattern cluster.

This completes the proof of Statement 2.

Proof of Statement 3. Consider the cluster v_a^{inv} obtained as a result of ordinal-invariant pattern clustering. We choose an arbitrary object x_i^* in this cluster and arrange its parameters in non-decreasing order: $x_{i1}^* \leq \ldots \leq x_{ij}^* \leq \ldots \leq x_{in}^*$. Graphically, this means that the object is represented in parallel coordinates system as a non-decreasing polyline. We denote this sequence of the positions of parameters $(x_{i1}^*, \ldots, x_{ij}^*, \ldots, x_{in}^*)$ by $P^{'}$.

Due to the order invariance of the pattern cluster v_a^{inv}, polylines of all the objects that occur in it have the same form for any sequence of indices, including the sequence $P^{'}$. Therefore, for a sequence of parameters $P^{'}$ all objects are represented parallel coordinates system as non-decreasing polylines.

This completes the proof of Statement 3.

Proof of Statement 4. We use the theorem proved in [15]: "Two objects x_1 and x_2 defined by vectors $x_1 = (x_{11}, \ldots, x_{1j}, \ldots, x_{1n})$ and $x_2 = (x_{21}, \ldots, x_{2j}, \ldots, x_{2n})$ respectively belong to the same ordinal-invariant pattern cluster if and only if they can be represented by complete weighted digraphs G_1 and G_2 with identical weights on the edges... that connect their respective vertices". The values of the edges are determined by formula (4), i.e., by the values of pairwise comparisons of the corresponding vertices.

For source objects x_1 and x_2, we construct the corresponding digraphs G_1 and G_2, as well as the digraph G_s of the resulting object $x_s = x_1 + x_2$. Since, according to the condition, these objects belong to the same ordinal-invariant pattern cluster, the values of the corresponding edges (pairwise comparisons) of digraphs G_1 and G_2 are the same. We need to show that the values of the edges of the digraph G_s also coincide with them. This follows from the following already established properties:

$$\begin{cases} x_{ij} > x_{ik} \\ x_{zj} > x_{zk} \end{cases} \Rightarrow (x_{ij} + x_{zj}) > (x_{ik} + x_{zk}),$$
$$\begin{cases} x_{ij} = x_{ik} \\ x_{zj} = x_{zk} \end{cases} \Rightarrow (x_{ij} + x_{zj}) = (x_{ik} + x_{zk}),$$
$$\begin{cases} x_{ij} < x_{ik} \\ x_{zj} < x_{zk} \end{cases} \Rightarrow (x_{ij} + x_{zj}) < (x_{ik} + x_{zk}),$$

where $j, k = 1, \ldots, n$.

This implies that, first, encoding of the new object x_s will match the encodings of objects x_1 and x_2 and, second, values of the edges of digraph G_s will match the edge values of G_1 and G_2.

This completes the proof of Statement 4.

REFERENCES

1. Fisher, R.A., The Use of Multiple Measurements in Taxonomic Problems, *Ann. Eugenics*, 1936, vol. 7, pp. 179–188.

2. Shawe-Taylor, J. and Cristianini, N., *Kernel Methods for Pattern Analysis*, Cambridge: Cambridge Univ. Press, 2004.

3. Duda, R.O., Hart, P.E., and Stork, D.G., *Pattern Classification*, New York: Wiley, 2012.

4. Mirkin, B.G., *Clustering for Data Mining: A Data Recovery Approach*, Boca Raton–London–New York–Singapore: Chapman & Hall/CRC, 2005.

5. Mirkin, B.G., Summary and Semi-Average Similarity Criteria for Individual Clusters, in *Models, Algorithms, and Technologies for Network Analysis*, New York: Springer, 2013, vol. 59, pp. 101–126.

6. Aleskerov, F.T., Solodkov, V.M., and Chelnokova, D.S., Dynamical Analysis of Behaviour Patterns for Russian Commercial Banks, *Ekonom. Zh. Vyssh. Shk. Ekonom.*, 2006, vol. 10, no. 1, pp. 48–62.

7. Aleskerov, F., Ersel H., Yolalan R., Clustering Turkish Commercial Banks According to Structural Similarities, *Yapi Credit Bank, Discussion Paper Ser.*, 1997, no. 97–02.

8. Cole, R. and Gunther, J., *A CAMEL Rating's Shelf Life*, Federal Reserve Bank of Dallas, December, 1995.

9. Gardner, M. and Mills, D.L., *Managing Financial Institutions: An Asset. Liability Approach*, Fort Worth: Dryden Press, 1994.

10. Golin, J., *The Bank Credit Analysis Handbook: A Guide for Analysts, Bankers and Investors*, New York: Wiley, 2001.

11. Aleskerov, F., Ersel, H., Gundes, C., and Yolalan, R., Multicriterial Method for Personnel Allocation among Bank Branches, *Yapi Kredi Discus. Paper Ser.*, Istanbul, 1998, no. 98–01.

12. Aleskerov, F.T. and Alper, C.E., Inflation, Money, and Output Growths: Some Observations, *Bogazici Univ. Res. Paper*, 1996, no. SBE 96-06.

13. Aleskerov, F.T., Belousova, V.Yu., Egorova, L.G., and Mirkin, B.G., Pattern Analysis in Statics and Dynamics, Part I, *Biznes-Informatika*, 2013, vol. 3, pp. 3–18.

14. Few, S., *Multivariate Analysis Using Parallel Coordinates*, 2006. https://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf (accessed at 28.03.2018).

15. Myachin, A.L., Pattern Analysis: Ordinal-Invariant Pattern Clustering, *Upravlen. Bol'shimi Sist.*, 2016, no. 61, pp. 1–59.

16. Inselberg, A., *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*, Tokyo: Springer, 2009.

17. Inselberg, A., The Plane with Parallel Coordinates, *Visual Comput.*, 1985, vol. 1 (2), pp. 69–91.

18. Myachin, A.L., New Methods of Pattern Analysis in the Study of Iris Anderson-Fisher Data, *6 Int. Conf. Comput. Commun. Control* (ICCCC), 2016, pp. 97–102.

19. Aleskerov, F.T., Egorova, L.G., Gokhberg, L.M., et al., Pattern Analysis in the Study of Science, Education and Innovation Activity in Russian Regions, *Procedia Comput. Sci.*, 2013, vol. 17, pp. 687–694.

20. *Reiting innovatsionnogo razvitiya sub"ektov Rossiiskoi Federatsii: analiticheskii doklad* (Rating of Innovation Development of Russian Federation Regions), Gokhberg, L.M., Ed., Moscow: Nats. Issled. Univ. "Vysshaya Shkola Ekonomiki," 2012.

This paper was recommended for publication by A.I. Mikhal'skii, a member of the Editorial Board