Mathematical Analysis of Synthetic Measures Based on Radar Charts

Boleslaw Borkowski, Artur Wiliński, Wieslaw Szczesny and Zbigniew Binderman

Warsaw University of Life Sciences – SGGW

Nowoursynowska str. 159, 02-776 Warsaw, Poland

E-mail(*corresp.*): `artur.piotr.wilinski@gmail.com`

E-mail: `artur_wilinski@sggw.pl`

Abstract. This work contains a description of a technique for constructing two synthetic indicators (measures) using a graphical presentation in the form of radar maps. The paper presents the structure and properties of indicators and their formal notation specially created for this purpose using the analogon of a scalar product of vectors. In particular, it proves the theorem on polygon fields, induced by radar maps, prepared for structural vectors, which allows to build concentration indicators. In order to demonstrate the usefulness of tools constructed by such means, the example shows how significant structural changes can be imperceptible when utilizing only the GINI concentration indicator's value, but are noticeable when using the concentration indicator developed by the authors. In addition, it illustrates the change in the value of concentration indicators (GINI and the indicator developed by the authors) on two families of Lorenz curves, together with changes in concentration. The practical application of this technique for constructing indicators that create rankings is presented on empirical data on the level of material deprivation in the countries that joined the EU in 2004 and 2007. These data have also been annotated (for comparison purposes) using the so-called overrepresentation maps (Grade Correspondence Analysis method).

Keywords: radar chart, concentration measure, **S**-shift operator.

AMS Subject Classification: 65H10; 68W30.

1 Introduction

The graphical presentation of data plays a very important role in the Multidimensional Data Analysis. The method of data presentation leads to the con-

struction of measures/indicators for evaluation the objects described by these data, in particular, to organize them and build various data rankings (as for example in [12, 13]). However, apart from graphical interpretation, the basic aspect of each measurement is its basic properties. Sometimes these methods give us misconceptions about data structure. Radar charts are a useful tool for graphical analysis of multidimensional data [11]. In the literature, the name *radar charts* is also known under other names [11]: *polar graph*, *Kiviata diagram*, *star chart*, *irregular polygon*, etc.

Let us assume that the variables $X_1, ..., X_n$, with non-negative values, describing different objects, have identical weights and are expressed on the same scale. Then, for the selected object $\mathbf{x} = (x_1, ..., x_n)$, the value of the synthetic index W of the x object can be taken as the chart field, with the bases of particular bars equal to $\frac{1}{n}$ and values $x_1, ..., x_n$. Let the vector $x_1, ..., x_n$, where $0 \leq x_1 \leq x_2 \leq ... \leq x_n$ and $x_1 + ... + x_n = 1$, show a specific structure for a certain population of $Z_s, (s = 1, ..., m)$ composed of m_s objects. However, let the vector $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_n)$, where for $i = 1, ..., n$, denotes the accumulation of the structure. The *GINI* concentration factor, calculated for a given structure, is expressed as a supplement to the one of the quotient of fields between the OX axis and two curves determined by the sets of points: the first curve $\left((0;0), (\frac{1}{n}; \hat{x}_1), (\frac{2}{n}; \hat{x}_2), ..., (\frac{n-1}{n}; \hat{x}_{n-1}), (1;1)\right)$ and the second curve $\left((0;0), (\frac{1}{n}; \frac{1}{n}), (\frac{2}{n}; \frac{2}{n}), ..., (\frac{n-1}{n}; \frac{n-1}{n}), (1;1)\right)$ (see in [6, 7]).

Graphically, these curves are the diagonal of the unit square and the Lorenz curve (see in [1]). Both of these, synthetic index W and *GINI* measures, are used as tools for multidimensional comparative analysis, based on graphs of objects in the Cartesian system. Using radar charts, we can create the equivalents of these two indicators [2, 3, 4]. In order to determine the properties of these indicators, the authors defined analogs of the scalar product of vectors as **S**-scalar product) and the standard of the vector (under the name **S**-norm). The square of the vectors standard considered here is equal (to the nearest constant multiplier) to the field of the polygon formed by the radar plot of these vectors.

The work presents the properties of two measures, which in practical applications can be used to build rankings of objects described by many variables and to assess the level of concentration of features $\mathbf{f}$ or objects $\mathbf{o}$ in a set (population) (see in [13, 14]). Radar methods have proved very useful for data analysis in scientific research (e.g. see in [2]) and e.g. Big data (see in [12]) due to the easy visualization of multidimensional data. These methods meet the basic postulate of stability of the applied method, for example they do not depend on the ordering of the vectors features describing the given object.

It is worth noting that in addition to the definition of the indicator itself using geometric interpretation, a very important issue is to determine the properties of these indicators. For this purpose we have attempted to describe them formally. Some properties of the indicators discussed in earlier papers (see [4] and [3]). In this paper, an analogon of the scalar product of vectors (under the name **S**-scalar product) and the vector norm (under the name **S**-norm) has been introduced.

These concepts were used to formally present the technique of creating and

study properties of indicators (using the idea of presentation in the form of radar charts). The usefulness of the developed indicators was presented by the Authors in this paper on empirical data on the level of material deprivation in the countries that joined the EU in 2004 and 2007. The presentation of data in the form of overrepresentation maps was also used for this purpose.

The concepts discussed in this work and our previous works were used to formally present the technique for creating and representing the properties of indicators using radar maps. This work presents for the first time the formal construction of an analogon of a scalar product of vectors used to build indicators and study their properties. The work consists of 5 sections and an abstract, conclusions and bibliography. Section 2 presents a mathematical description of building synthetic indicators with the use of radar maps and a verification of their properties. Section 3 presents the formal structure of the construction of a synthetic indicator for ordering objects. Section 4 shows how a concentration indicator can be created and performs a comparative analysis of the properties of the two discussed indicators (GINI and the one constructed by the authors). Finally, in the last Section 5 we discuss the results of the analysis of empirical data. In addition, we performed a similar analysis of empirical material using overrepresentation maps in order to compare the two methods.

2   S-shift operator

2.1   S-scalar product

Let us consider here and define the $\mathbf{S}$-shift operator. Let $\mathbf{X} = \mathbb{R}^n$ be the n-dimensional space of Euclid. The set of all linear operators whose fields are equal ($\mathbf{X}$ space with linear subsets) and whose values also belong to the $\mathbf{X}$ space is denoted by $L_0(\mathbf{X})$.

DEFINITION 1. $\mathbf{S}$-operator $\mathbf{S} : \mathbf{X} \to \mathbf{X}$ will be the $\mathbf{S}$-shift operator if for everyone $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbf{X}$:

$$\mathbf{Sx} := (x_2, x_3, ..., x_n, x_1).$$

Theorem 1. *Let* $\mathbf{I}$ *be the identifier operator. The* $\mathbf{S}$*-operator has the following properties: a)* $\mathbf{S} \in \mathbf{L_0(X)}$*, b)* $\mathbf{S}^n = \mathbf{I}$*, c)* $\mathbf{S}$*-operator is a reversible operator, i.e. there is such an operator* $\mathbf{S}^{-1} \in \mathbf{L_0(X)}$ *that we have on* $\mathbf{X}$ *:* $\mathbf{SS}^{-1} = \mathbf{S}^{-1}\mathbf{S} = \mathbf{I}$*, d) if* $\mathbf{a} \in \mathbf{X}$ *is a constant vector, that is* $\mathbf{a} = (a, a, ..., a)$ *then* $\mathbf{Sa} = \mathbf{a}$*.*

Proof.   a). Let $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbf{X}$ and $\mathbf{y} = (y_1, y_2, ..., y_n) \in \mathbf{X}$, $\alpha, \beta \in \mathbb{R}$ be freely set vectors, numbers, respectively. By definition, we have:

$$\begin{aligned}
\mathbf{S}(\alpha\mathbf{x} + \beta\mathbf{y}) &= \mathbf{S}(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, ..., \alpha x_n + \beta y_n) \\
&= (\alpha x_2 + \beta y_2, \alpha x_3 + \beta y_3, ..., \alpha x_n + \beta y_n, \alpha x_1 + \beta, y_1) \\
&= \alpha(x_2, x_3, ..., x_n, x_1) + \beta(y_2, y_3, ..., y_n, y_1) = \alpha\mathbf{Sx} + \beta\mathbf{Sy}.
\end{aligned}$$

c). The reverse operator to $\mathbf{S}$ is $\mathbf{S}^{-1}\mathbf{x} := (x_n, x_1, x_2, ..., x_{n-1})$, where $\mathbf{x} = (x_1, x_2, ..., x_n)$. It's easy to check that $\mathbf{SS}^{-1} = \mathbf{S}^{-1}\mathbf{S} = \mathbf{I}$.

Statements *b*) and *d*) follow directly from definition.   □

DEFINITION 2. The mapping of the product of the space $\mathbf{X} \times \mathbf{X}$ into a set of real numbers $\mathbb{R}$, determined by the formula: $\langle \mathbf{x}, \mathbf{y} \rangle_S := \langle \mathbf{x}, \mathbf{Sy} \rangle$ we call $\mathbf{S}$-scalar product of $\mathbf{x}$ and $\mathbf{y}$ vectors, where the symbol $\langle \cdot \rangle$ means the scalar product. It is easy to check that the $\mathbf{S}$-scalar product so defined has the following properties:

Theorem 2. *If* $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ *are given vectors and* $\alpha, \beta \in \mathbb{R}$ *then:*

a) $\langle \alpha\mathbf{x}, \mathbf{y} \rangle_S = \langle \mathbf{x}, \alpha\mathbf{y} \rangle_S = \alpha\langle \mathbf{x}, \mathbf{y} \rangle_S,$ *b)* $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle_S = \langle \mathbf{x}, \mathbf{z} \rangle_S + \langle \mathbf{y}, \mathbf{z} \rangle_S,$

c) $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle_S = \langle \mathbf{x}, \mathbf{y} \rangle_S + \langle \mathbf{x}, \mathbf{z} \rangle_S,$ *d)* $\langle \mathbf{x}, \mathbf{x} \rangle_S = \langle \mathbf{Sx}, \mathbf{x} \rangle,$ *e)* $\langle \mathbf{Sx}, \mathbf{Sy} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle.$

Note that the scalar $\mathbf{S}$-scalar product does not satisfy the third and fourth axioms of the scalar product of linear space elements (see in [5]). The following example proves this.

Example 1. Let $n=3$, $\mathbf{x} = (1,0,0)$, $\mathbf{y} = (0,0,1)$, the symbol of $\circ$ means the scalar multiplication operation. We have then:

$$\mathbf{Sx} = \mathbf{y}, \ \mathbf{Sy} = (0,1,0), \ \langle \mathbf{x}, \mathbf{y} \rangle_S = \langle \mathbf{x}, \mathbf{Sy} \rangle = (1,0,0) \circ (0,1,0) = 0,$$
$$\langle \mathbf{y}, \mathbf{x} \rangle_S = \langle \mathbf{y}, \mathbf{Sx} \rangle = (0,0,1) \circ (0,0,1) = 1,$$
$$\langle \mathbf{x}, \mathbf{x} \rangle_S = \langle \mathbf{x}, \mathbf{Sx} \rangle = (1,0,0) \circ (0,0,1) = 0.$$

It follows from the above that $\langle \mathbf{x}, \mathbf{y} \rangle_S \neq \langle \mathbf{y}, \mathbf{x} \rangle_S$ and $\langle \mathbf{x}, \mathbf{x} \rangle_S = 0$ for the vector $\mathbf{x} \neq (0, ..., 0)$.

2.2   Pseudonorms generated by S-scalar product

Let us assume the following definition of a norm generated by a $\mathbf{S}$-scalar product.

DEFINITION 3. The function $\| \| \|_S : \mathbf{X} \to \mathbb{R}_+$, where $\mathbb{R}_+ := \{\alpha : \alpha \geq 0\}$ determined by the formula:

$$\|\mathbf{x}\|_S := |<\mathbf{x}, \mathbf{x}>_S|^{1/2} = \left(\frac{1}{n} | \sum_{k=1}^{n} x_k x_{k+1} | \right)^{1/2},$$

where $\mathbf{x} = (x_1, x_2, ..., x_n), x_{n+1} := x_1$ we will be called the $\mathbf{S}$-norm and the number $\|\mathbf{x}\|_S$ the $\mathbf{S}$-norm of the $\mathbf{x}$-vector. In some cases the $\mathbf{S}$-norm can be given a simple geometrical interpretation.

Example 2. Let $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, $1 \geq x_i \geq 0, i = \{1, 2, ..., n\}, n \geq 2$. On the plane, we enter the regular n-polygon into a unit circle (with radius $r=1$) with the center at the beginning of the coordinate system and connect the vertices of this polygon with the center of the system. The straight segments of length 1 obtained in this way are determined successively by $\mathbf{O}_1, \mathbf{O}_2, ..., \mathbf{O}_n$ (to draw attention), starting from the section lying on the vertical axis. Co-ordinates of the vector $\mathbf{x}$ can be represented by means of a radar plot. For this purpose, let us denote by $\mathbf{x}_i$ the points on the axis 0_i arising from the intersection of the axis 0_i with the circle with the center at the beginning of the system and the radius equal to $x_i, i = 1, 2, ..., n$.
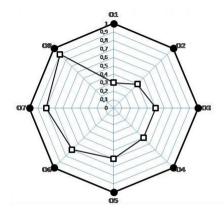
Figure 1. Radar chart for vector $\mathbf{x} = (0.3, 0.4, 0.5, 0.5, 0.6, 0.7, 0.8, 0.9)$. Painted circles are the tops of a regular octagon, open squares are the coordinates of the x vector. Source: own study.

The following Figure 1 (radar chart) gives the illustrations for the vector: $\mathbf{x} = (0.3, 0.4, 0.5, 0.5, 0.6, 0.7, 0.8, 0.9)$. Painted dots indicate the vertices of a regular polygon (octagon). The empty squares correspond to the coordinates of the vector $\mathbf{x}$.

By connecting the points in order x_1 with x_2, x_2 with x_3, ..., x_n with x_1 we obtain a n-polygon ($n = 8$) whose field $\mathbf{S}_1$ (area of the polygon defined by the vector $\mathbf{x}$ – points marked with the squares) is defined by the formula:

$$\mathbf{S}_1 = \sum_{i=1}^{n} \frac{1}{2} x_i x_{i+1} \sin\left(\frac{2\pi}{n}\right) = \frac{1}{2} \sin\left(\frac{2\pi}{n}\right) \sum_{i=1}^{n} x_i x_{i+1} = \frac{1}{2} n \sin\left(\frac{2\pi}{n}\right) \|\mathbf{x}\|_S^2,$$

where $x_{n+1} := x_1$. The area of the regular n-polygon inscribed in a circle (of radius 1) specifies the formula:

$$\mathbf{S}_0 = \sum_{i=1}^{n} \frac{1}{2} \cdot 1 \cdot 1 \cdot \sin\left(\frac{2\pi}{n}\right) = \frac{1}{2} n \sin\left(\frac{2\pi}{n}\right),$$

whereas the ratio of fields of these polygons $\mathbf{S}_1/\mathbf{S}_0$ is determined by the number:

$$\hat{\mathbf{S}}(\mathbf{x}) := \frac{\mathbf{S}_1}{\mathbf{S}_0} = \frac{1}{n} \sum_{i=1}^{n} x_i x_{i+1} = \|\mathbf{X}\|_S^2,$$

From this follows the following note.

Attention. The square of the **S**-norm of the vector $\mathbf{x}$ is the quotient of the area of the polygon induced by this vector and the field of the regular n-polygon, induced by the unit vector, i.e. $\|\mathbf{X}\|_S^2 = \hat{\mathbf{S}}$. So the question is whether $W(\mathbf{x}) := \|x\|_S$ can be a good indicator for organizing linear objects (each of which is described with n values). The considerations presented below allow us to resolve this dilemma.

2.3   S-norms of vectors belonging to the cone $\mathbb{R}^n_+$

Let be the cone in space $\mathbf{R}^n$

$$\mathbf{x} = \mathbb{R}^n_+ := \{\mathbf{x} = (x_1, x_2, \cdots, x_n) : x_i \geq 0, i = 1, 2, \cdots n\},$$

vectors $\mathbf{0} = (0, 0, \cdots, 0)$, $\mathbf{1} = (1, 1, \cdots, 1)$. Let

$$\mathbf{x} = (x_1, x_2, \cdots, x_n), \mathbf{y} = (y_1, y_2, \cdots, y_n) \in \mathbf{X}.$$

If

$$x_k > y_k, (x_k \geq y_k), \text{for } k = 1, 2, \cdots, n,$$

then we will write that $\mathbf{x} > \mathbf{y}$, $(\mathbf{x} \geq \mathbf{y})$.

It is not difficult to notice that if $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ it is natural to call the vector $\mathbf{x}$ better (higher evaluated) than the vector $\mathbf{y}$. This means that none of the components of the vector $\mathbf{x}$ is smaller than the corresponding components of the vector $\mathbf{y}$, and at least one of them has a higher value, i.e. there is such a $k \in [1, n]$ that $x_k > y_k$.

Let the equality $h(\mathbf{x}) = \langle \mathbf{x}, \mathbf{x} \rangle_S$ be satisfied. Assume the following designations for the mapping kernel h: $kerh := \{\mathbf{x} \geq \mathbf{0} : h(\mathbf{x}) = 0\}$, and

$$\mathbf{N}_e := \{\mathbf{x} = (0, x_2, 0, x_4, 0, \cdots, 0, x_n) \vee \mathbf{x} = (x_1, 0, x_3, 0, x_5, 0, \cdots, x_{n-1}, 0)\},$$

when n is an even number,

$$\mathbf{N}_o := \{\mathbf{x} = (0, x_2, 0, x_4, 0, \cdots, 0, x_n) \vee \mathbf{x} = (x_1, 0, x_3, 0, \cdots, 0, x_{n-2}, 0, 0)\},$$

when n is an odd number and

$$\mathbf{N} := \begin{cases} \mathbf{N}_e, & \text{when } n \text{ is an even number,} \\ \mathbf{N}_o, & \text{when } n \text{ is an odd number.} \end{cases}$$

The following theorem can be proved.

Theorem 3. *Scalar product $\langle \mathbf{x}, \mathbf{x} \rangle_S = \mathbf{0} \Leftrightarrow \mathbf{x} \in \mathbf{N}$.*

Conclusion. Let it be $\mathbf{x} \in \mathbf{X}$, S-norm of vector $\|\mathbf{x}\|_S = 0 \Leftrightarrow \mathbf{x} \in \mathbf{N}$. If the $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ vectors meet the conditions of $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$, then it is obvious that $\|\mathbf{x}\|_S \geq \|\mathbf{y}\|_S$.

For $\mathbf{x}, \mathbf{y} \in \mathbf{N} : \mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ equality is fulfilled $\|\mathbf{x}\|_S = \|\mathbf{y}\|_S$. The equality of these norms can also occur for $\mathbf{x}, \mathbf{y} \notin \mathbf{N}$. This is illustrated by the following Example 3.

Example 3. Let $\mathbf{x} = (1, 0, 2, 0, 1)$, $\mathbf{y} = (1, 0, 1, 0, 1)$. Then we have that $\mathbf{x}, \mathbf{y} \notin \mathbf{N}_0, \mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$,

$$\mathbf{Sx} = (0, 2, 0, 1, 1), \quad \mathbf{Sy} = (0, 1, 0, 1, 1),$$
$$(\|\mathbf{x}\|_S)^2 = (1 \cdot 0 + 0 \cdot 2 + 2 \cdot 0 + 0 \cdot 1 + 1 \cdot 1)/5 = 1/5,$$
$$(\|\mathbf{y}\|_S)^2 = (1 \cdot 0 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1)/5 = 1/5.$$

From here we have $\|\mathbf{x}\|_S = \|\mathbf{y}\|_S$. The following lemma specifies for which $\mathbf{x}, \mathbf{y} \in \mathbf{X} \setminus \mathbf{N}$, the inequality of $\|\mathbf{x}\|_S > \|\mathbf{y}\|_S$ is met. The following property is true.

Property 1. If the vectors $\mathbf{x} = (x_1, x_2, ..., x_n), \mathbf{y} = (y_1, y_2, ..., y_n) \in \mathbf{X}$ meet the conditions $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{y} \neq \mathbf{x}$ and there is such a $k \in [1, n] : x_k > y_k$ and $x_{k-1}x_{k+1} > 0$, then $\|\mathbf{x}\|_S > \|\mathbf{y}\|_S$.

The following conclusion follows from the above property.

Conclusion. If $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ vectors meet the conditions $\mathbf{x} \geq \mathbf{y} > 0$ and $\mathbf{x} \neq \mathbf{y}$ then $\|\mathbf{x}\|_S > \|\mathbf{y}\|_S$.

The following theorem can be proved.

Theorem 4. *Function $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n) := \|\mathbf{x}\|_S, \mathbf{x} \in \mathbf{X}, \mathbf{x} > 0$, with fixed values of $x_1, x_2, \cdots, x_{j-1}, x_{j+1}, \ldots, x_n$ is an increasing function of the real variable $x_j > 0$, for $j \in [1, n]$.*

Let's take the following note:

$$\mathbf{N}_1 := \{\mathbf{x} \in \mathbf{X} : \mathbf{x} = (0, 0, \ldots, 0, \mathbf{x}_i, 0, \ldots, 0, 0), i \in [1, n]\},$$

i.e. set $\mathbf{N}_1$ contains all those vectors $\mathbf{x} \geq 0$, for which all coordinates - except perhaps one, are equal to zero. It's easy to see that $\mathbf{N}_1 \subset \mathbf{N}$. Let the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{X}$ be arbitrarily set. Let us denote the j-th permutation of the set of $\mathbf{x}$-coordinates by $\mathbf{x}_j := (x_{1j}, x_{2j}, \cdots, x_{nj})$, where $j = 1, 2, ..., n!, \mathbf{x}_1 := \mathbf{x}$.

You can see that the following note is true.

Attention. If $\mathbf{x} \in \mathbf{N}_1$ then $\|\mathbf{x}\|_S = 0$ for each $j = 1, 2, \ldots, n!$

Further, without losing the general nature of our considerations, one can make the following assumption that regarding the permutation of the set of coordinates of the considered vectors: if $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ satisfy the conditions: $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ then these conditions are also met by vectors that are permutations of their coordinates, i.e. $\mathbf{x}_i \geq \mathbf{y}_i$ and $\mathbf{x}_i \neq \mathbf{y}_i$ for each $j = 1, 2, \ldots, n!$

The following statements can be proved.

Theorem 5. *If the vector $\mathbf{x} \in \mathbf{X} \backslash \mathbf{N}_1$ then there is permutation of the vector $\mathbf{x}$-vector $\mathbf{x}_j$ coordinates, which $\mathbf{S}$-norm is greater than zero, i.e. there is such a $j \in \{1, 2, \cdots, n!\}$ that $\|\mathbf{x}_j\|_S > \mathbf{0}$.*

Theorem 6. *If the vectors $\mathbf{x} \in \mathbf{X} \backslash \mathbf{N}_1$ and $\mathbf{y} \in \mathbf{X}$ meet the conditions: $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ then there are such permutations $\mathbf{x}_j, \mathbf{y}_j$ vectors $\mathbf{x}, \mathbf{y}$ that $\|\mathbf{x}\|_S > \|\mathbf{y}\|_S$, where $j \in \{1, 2, \cdots, n!\}$.*

Theorem 7. *Let $n > 1$ and vector $\mathbf{x} \in \mathbf{X} = \mathbb{R}^n_+$, natural number $k := n!, \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k \in \mathbf{X}$ be the vectors obtained from the vector $\mathbf{x}$ through all permutations of its coordinates. If $\mathbf{x} \in \mathbf{X} \backslash \mathbf{N}_1$ then a finite set of k-numbers of $A := \{\|\mathbf{x}_1\|_S, \|\mathbf{x}_2\|_S, \cdots, \|\mathbf{x}_k\|_S\}$ contains at most $0.5(n-1)$ different values.*

Next, we will formulate further definition.

DEFINITION 4. If $\mathbf{x} \in \mathbf{X}$ then the numbers:

$$\|\mathbf{x}\|_M := \max_{1 \leq j \leq k} \|\mathbf{x}_j\|_S, \quad \|\mathbf{x}\|_\Xi := \frac{1}{k} \sum_{j=1}^{k} \|\mathbf{x}_j\|_S, \quad \|\mathbf{x}\|_m := \min_{1 \leq j \leq k} \|\mathbf{x}_j\|_S$$

will be called the maximum, average and minimum norm of the $\mathbf{x}$ vector, respectively.

Directly from the definition it follows that:
Attention. If the vector $\mathbf{x} \in N_1$, then $\|\mathbf{x}\|_M = \|\mathbf{x}\|_S = \|\mathbf{x}\|_m = 0$.
Attention. For each $\mathbf{0} \neq \mathbf{x} \in \mathbf{X} : \|\mathbf{x}\|_M \geq \|\mathbf{x}\|_\Xi \geq \|\mathbf{x}\|_m > 0$.

3   Synthetic measures induced by S-norms

Let the $\mathbf{x} \in \mathbf{X} = \mathbb{R}^n_+$ vector, the natural number $k := n!$, $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k \in \mathbf{X}$ be the vectors obtained from the vector $\mathbf{x}$ through all permutations of its coordinates. The following definition is met.

DEFINITION 5. Let it be a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \mathbf{X}$ and $\bar{x} = (x_1 + x_2 + \ldots + x_n)/n$. The numbers

$$M(\mathbf{x}) := \begin{cases} \|\mathbf{x}\|_M & \text{for } x \in X \backslash N_1, \\ \bar{x} & \text{for } x \in N_1, \end{cases}$$

$$s(\mathbf{x}) := \begin{cases} \|\mathbf{x}\|_\Xi & \text{for } x \in X \backslash N_1, \\ \bar{x} & \text{for } x \in N_1, \end{cases} \qquad m(\mathbf{x}) := \begin{cases} \|\mathbf{x}\|_m & \text{for } x \in X \backslash N_1, \\ \bar{x} & \text{for } x \in N_1 \end{cases}$$

will be called *radar synthetic measures* of the vector $\mathbf{x} \in \mathbf{X}$: maximum, medium and minimum, respectively.

From the definitions and previous properties follows the following note.
Attention. If $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ meet the conditions: $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ then

$$a)\, M(\mathbf{x}) > M(\mathbf{y}), \quad b)\, S(\mathbf{x}) > S(\mathbf{y}), \quad c)\, m(\mathbf{x}) \geq m(\mathbf{y}).$$

The following theorems follow.

Theorem 8. *Let* $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{x} \in \mathbf{X} \backslash \mathbf{N}_1$ *then at fixed values* $x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ *radar synthetic measures* $(m(\mathbf{x}), s(\mathbf{x}), M(\mathbf{x}))$ *are functions that grow in the real variable* $x_j \geq 0$, *for* $j \in [1, n]$.

Theorem 9. *If* $\mathbf{x}$, $\mathbf{a} = (a, a, \ldots, a) \in \mathbf{X}, \alpha \in \mathbb{R}_+$, $\mathbf{1} = (1, \ldots, 1)$, *then:*

a. $M(\alpha \mathbf{x}) = \alpha M(\mathbf{x}); S(\alpha \mathbf{x}) = \alpha S(\mathbf{x}); m(\alpha \mathbf{x}) = \alpha m(\mathbf{x})$,

b. $M(\mathbf{a}) = M(a\mathbf{1}) = aM(\mathbf{1}) = a$;
$S(\mathbf{a}) = S(a\mathbf{1}) = aS(\mathbf{1}) = a$;
$m(\mathbf{a}) = m(a\mathbf{1}) = am(\mathbf{1}) = a$,

c. *If* $\mathbf{0} \leq \mathbf{x} = (x_1, x_2, \ldots, x_n) \leq \mathbf{1}$, *then*
$0 \leq m(\mathbf{x}) \leq S(\mathbf{x}) \leq M(\mathbf{x}) \leq 1$.

It follows from the above considerations that the indicator $W(\mathbf{x}) := \|\mathbf{x}\|_S$ can be used only in special practical situations, because it has undesirable properties in general (its value depends on the order of the $\mathbf{x}$-coordinates and assumes a value of zero in too many cases).

A typical example of its use would be to compare the concentration of features (good) in different populations, which will be described in the next section. On the other hand, the meters $m(\mathbf{x}), S(\mathbf{x})$ and $M(\mathbf{x})$ no longer have these disadvantages for $\mathbf{x} \in \mathbf{X} \backslash \mathbf{N}_1$.

4   Measure of concentration induced by S-norm

4.1   Concentration measurement

Practitioners studying the degree of income and resource diversification in the possession of a certain group of objects most often present the following postulates regarding the $d(\mathbf{x}, \mathbf{y})$ indicator used for measurement (in the case of aggregated data it concerns the differentiation between the structure of $\mathbf{x}$ units and the structure of $\mathbf{y}$ resources owned by them):

I.   the index assumes a value equal to zero if the good is distributed evenly among all objects (the both structures are identical, i.e. $\mathbf{x} = \mathbf{y}$);

II.   the values of the indicator are consistent with the principle of transfers, which states that the transfer of a poorer object to any part of its resources to the richer always entails an increase in inequality in the population (this means that the flow of values between the components of x_i and x_{i+s} increases the value of $d(\mathbf{x}, \mathbf{y})$);

III.   sensitivity of transfer (transfer sensitivity axiom): the impact on transfer of goods from a "poor" object to a "poorer" object to a change in the value of the indicator, with a constant transfer rate, the higher the value of good in possession of the object from which the transfer was made (i.e. further away is the component from which the transfer originates in relation to the component to which the transfer affects the bigger the value should change);

IV.   the value of the indicator will not change if the same good proportions in the possession of particular objects (such as scale invariance axiom) change with the same proportional changes.

The most popular indicator used to measure the concentration (inequality) of the distribution that meets these demands is the *GINI* coefficient defined as the doubled field between the *Lorenz* curve and the diagonal of the unit square (see in [1,8]). The next section will present the method of creating the measurement using the presentation of good structures in the form of radar charts.

4.2   Concentration measure induced by a radar vector graph

Let be

$$\mathbf{X} \supset \mathbb{R}_+^n := \{\mathbf{x} = (x_1, x_2, \ldots, x_n) : x_i \geq 0, i = 1, 2, \ldots, n\}, n \in \mathbf{N},$$

$$\mathbb{R}_+^n \supset \Omega := \{\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}_+^n : \sum_{i=1}^{n} x_i = 1\}.$$

Elements of $\mathbf{\Omega}$ set are called structural vectors or short structures. The vector $\mathbf{e} = (1/n, 1/n, \cdots, 1/n)$ will be called an egalitarian structure and vector $\mathbf{o} = (0, 0, \cdots, 0, 1)$ will be called an extremely concentrated structure.

Let be $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{X}$, the set $A := \{x_1, x_2, \ldots, x_n\}$. The vector $\mathbf{x}' := (x'_1, x'_2, \ldots, x'_n)$ will be called the ordered vector, where the following order $x'_1 \leq x'_2 \leq \ldots \leq x'_n$ is the permutation of the set A and vector $\mathbf{x}^\frown := (x_1^\frown, x_2^\frown, \ldots, x_n^\frown)$, for $x_i^\frown := \sum_{j=1}^{i} x_j, i = 1, 2, \ldots, n$ will be called the accumulation of the $\mathbf{x}$ vector.

If $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \Omega$ then by the symbol $\mathbf{x}'^\frown := (x_1'^\frown, x_2'^\frown, \ldots, x_n'^\frown)$, where $x_i'^\frown := \sum_{j=1}^{i} x'_j, i = 1, 2, \ldots, n$. we will understand the vector created respectively by ordering and accumulation of the $\mathbf{x}$-structure.

In order to define a concentration coefficient that uses radar vectors and has identical properties as the *GINI* coefficient, the following theorem can be proved.

Theorem 10. *If radar polygons $W_{\mathbf{x}'^\frown}$, $W_{\mathbf{e}^\frown}$ they are generated by $\mathbf{x}'^\frown$ and $\mathbf{e}'^\frown$ vectors, respectively, then $W_{\mathbf{x}'^\frown} \subseteq W_{\mathbf{e}'^\frown}$.*

Using the above theorem, it can be shown that the measurement based on the structure presentation in the form of a radar plot can be used to measure concentration. The following theorem can be proved.

Theorem 11. *If $P(\mathbf{x})$ denotes the area of the radar polygon induced by the structure $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \Omega$, then the coefficient*

$$G_R = 1 - \frac{P(\mathbf{x}'^\frown)}{P(\mathbf{e}^\frown)} = 1 - \frac{6n}{2n^2 + 4}\Big[\sum_{i=1}^{n-1} x_i'^\frown x_{i+1}'^\frown + x'_1\Big]$$

$$= 1 - \frac{6n^2}{2n^2 + 4}\|\mathbf{x}'^\frown\|_S^2,$$

where $x'^\frown := \sum_{j=1}^{i} x'_j$ it meets all four basic demands (I, II, III, IV) of the concentration.

DEFINITION 6. The measurement defined by the above formula will be called a radar concentration factor (e.g. income inequality). It can be shown directly from the definition that the G_R meets the conditions:

$$G_R(\mathbf{e}) = 0, \quad G_R(\mathbf{0}) = 1.$$

The obtained formula for G_R coefficient is similar to the formula for the standardized *GINI* coefficient (that is, the classic *GINI* measurement multiplied by $n/(n-1)$, then it meets the condition $G(\mathbf{o}) = 1$, which takes the form:

$$G^* = 1 - \frac{2}{n-1}\sum_{i=1}^{n-1} x_i'^\frown.$$

4.3   *GINI* concentration index vs G_R index

Let's consider two examples to illustrate the differences in response to changes in G_R and *GINI* indicators.
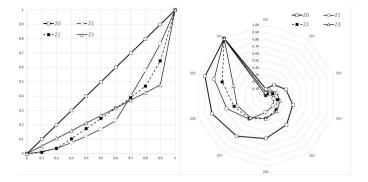
Figure 2. Left side: an exemplary data structure, the concentration of which is described by the Lorenz curves. Right side: there are clear differences between these curves for *GINI* and G_R indicators. Source: own study.

Example 4. We have ten objects, whose structures are presented in the rows of the Table 1, named $Z0, Z1, Z2, Z3$. The $Z0$ is an egalitarian structure. Graphically, the $Z0, \ldots, Z3$ structures are illustrated in Figure 2: on the left as Lorenz curves, and on the right they were presented as radar charts. There are clear differences between these curves for *GINI* and G_R indicators (see Figure 2). The polygon designated by $(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{n-1}, 1)$ is contained in the polygon designated by (see Figure 2) $(1/n, 2/n, \ldots, (n-1)/n, 1)$.

Table 1. The values of the *GINI* concentration index and the G_R index for the structures $(Z0 \div Z3)$ showing the distribution of income, where: *D01*, ..., *D10* are the names of the objects.

—	D01 [%]	D02 [%]	D03 [%]	D04 [%]	D05 [%]	D06 [%]	D07 [%]	D08 [%]	D09 [%]	D10 [%]	**GINI**	**GR**
Z0	10	10	10	10	10	10	10	10	10	10	**0**	**0**
Z1	1.30	2.30	4.00	4.40	5.02	6.00	17.00	17.80	18.80	23.38	**0.422**	**0.5242**
Z2	0.89	2.75	6.70	7.00	7.10	7.20	7.40	8.00	17.60	35.36	**0.422**	**0.5858**
Z3	5.31	5.31	5.31	5.31	5.31	5.31	5.32	5.32	5.32	52.18	**0.422**	**0.6446**

It's easy to check that the fields under Lorenz curves are identical, so *GINI* doesn't notice the changes, and the fields of the radar charts are getting smaller. Thus, the G_R indicator catches the differences between $Z0$ and the structures $Z1, Z2, Z3$.

The figure on the right side shows G_R measure and the structural differences between the curves (which can be seen on the radar charts – they differ in fields).

It illustrates how significant changes may be unnoticeable for the GINI indicator, but will be noted when we apply two *GINI* i and G_R indices (an index developed by the authors using radar charts) for the analysis. The level of material deprivation in Poland and Latvia in 2008, 2011 and 2017 is shown in Figure 6.

Example 5. The first example concerned specific changes between structures. In order to approximate the differences in sensitivity to changes in regular situations, we will present in the next example changes in 10-element structures that are built by discretization of the curves determined using the Lorenz curves

(see Figure 3). $L1(t) = t^a$, and $L2(t) = \Phi[\Phi^{-1}(t) - \mu], t \in [0,1]$ for

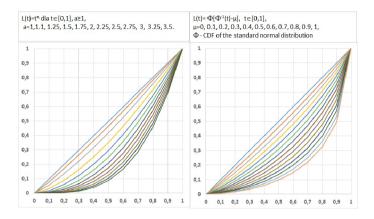$$\mu = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1].$$

Figure 3. Lorenz curves for different L(t). Source: own study.

The values of G_R and $GINI$ indicators are presented in Figure 4.
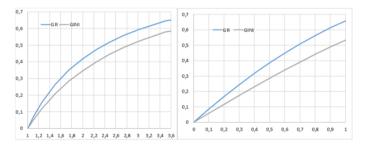

Figure 4. *GR* and *GINI* indicators. Source: own study.

Figure 3 presents Lorenz curves for structures obtained from curves defined by the formula:

$$(L_i(0.1), L_i(0.2) - L_i(0.1), \cdots, L_i(1) - L_i(0.9)),$$

$a = [1, 1.1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, .75, 3, 3.25, 3.5]$ and $\mu = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]$ for $i = 1$ (left side) and $i = 2$ (right side).

5   Data ranking due to the level of material deprivation

In this section, we will present the usefulness of indicators constructed on the basis of radar charts to present the ranking on the level of material deprivation in 12 countries that joined the EU in 2004 and 2017 (based on data from Eurostat).

Table 2.   Measure values $(\|\mathbf{x}\|_S)^2$ for 12 EU countries that joined the EU in 2004 and 2007. Source: own study based on EUROSTAT data.

| Countries | 2008 | | 2017 | | | |
	$\mathbf{W}_1 = \|\mathbf{X}\|^2$	$Rank_1$	$\mathbf{W}_2 = \|\mathbf{X}\|^2$	$Rank_2$	$\mathbf{W}_2 - \mathbf{W}_1$	$Rank$
Bulgaria	5.39	12	6.10	12	0.71	5
Czech Rep.	8.01	2	8.63	2	0.62	7
Estonia	8.19	1	8.33	4	0.14	10
Cyprus	7.57	5	7.38	8	-0.2	12
Latvia	6.72	9	7.32	9	0.6	8
Lithuania	7.50	6	7.32	10	-0.18	11
Hungary	6.61	10	7.49	7	0.88	4
Malta	7.76	4	8.67	1	0.91	2
Poland	6.92	8	8.11	5	1.19	1
Romania	5.77	11	6.67	11	0.88	3
Slovenia	8.00	3	8.35	3	0.35	9
Slovakia	7.30	7	7.95	6	0.65	6

Eurostat data can easily extract information on the population structure of individual 12 countries due to the declared number of material deprivation symptoms (for j-th country $\mathbf{x}_j = (100\%, 0\%, ..., 0\%)$ means that everyone in the country declares no deprivation symptoms). Material deprivation is the percentage of persons in households declaring the lack of possibility to satisfy several needs due to financial reasons (see in [9]). In other words, we have a vector

$$\mathbf{x}_j = (x_{j0}, x_{j1}, ..., x_{j9}), j = 1, 2, ..., 12$$

for each country, where x_{ji} for $i=0,1,...,9$ is the percentage of declared i-symptoms of material deprivation.

In this case, we have a naturally established order of coordinates of the vector $\mathbf{x}$ describing the structure of the population of a given country. The best situation is in a country in which it would be represented by the vector $\mathbf{x} = (1, 0, ..., 0)$, and the worst would be if the representation was realized by the vector $\mathbf{x} = (0, 0, ..., 0, 1)$.

It is easy to see that radar charts of $\hat{\mathbf{x}}_j$ vectors ($j=1,2,...,12$) are a good presentation, which in an intuitive way presents the situation regarding the level of material deprivation in such a way that the larger is the field of the polygon, the better is situation in the country. We present the visualization in the form of radar charts for 2017 in Figure 1, and the values of the meter $(\|(x)_S\|)^2$ for 2008 and 2017 in Table 2.

Table 2 (and Figure 5) shows that in 2017 the best situation is in Malta and the Czech Republic. However, compared to 2008, the situation has improved the most (in line with the change in the value of the measure) in Poland, Malta and Hungary, but it worsened in Cyprus and Lithuania.

Figure 6 illustrates the fact that in these two countries the changes between 2008 and 2011 went in opposite directions.

On this basis, radar maps are not very suitable for presenting data for many countries. A much better tool is the so-called overrepresentation maps (see Figure 7). For more on overrepresentation maps, see [9, 10].
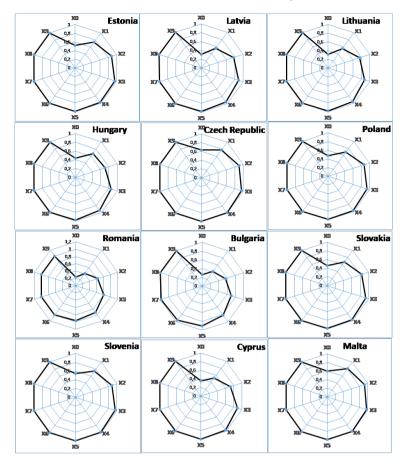
Figure 5. Presentation of the level of material deprivation in 2017. Source: own study based on EUROSTAT.

A typical example are presented radar charts and indicators using a polygon field. The presented considerations illustrate in a formal way the properties of this type of the measurement. If we order the descending structure of the 12 new EU member states according to this indicator in the period 2008–2017, then these structures can be presented in the form of an overrepresentation map (you can read more about them in the works (see in [9, 10]).

The degree of gray in this case presents the quotient of individual cells in a row to the average structure of these 12 countries for the period 2008–2017. As the average structure, we mean a vector, which is the arithmetic average of the row vectors. It is mean, that in each row the structure

$$(x_{i,0}, x_{i,1}, x_{i,2}, \ldots, x_{i,9}), x_{i,j} \geq 0$$

describes material deprivation. This structure is presented on the average structure

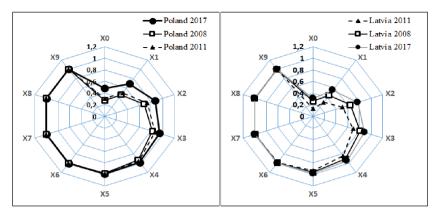$$(\overline{x}_0, \overline{x}_1, \overline{x}_2, \ldots, \overline{x}_{i,9})$$

Figure 6. Presentation of the level of material deprivation in 2008, 2011 and 2017. Source: own study based on EUROSTAT.
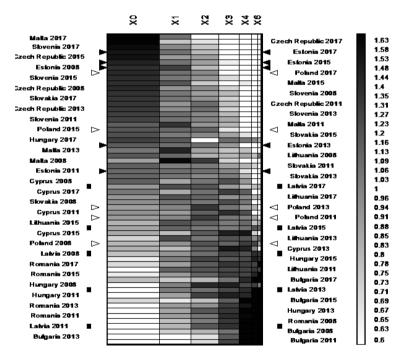

Figure 7. Overrepresentation map presenting the structures of the 12 new States due to the level of material deprivation. Source: own study.

of 12 EU countries in 2008,2011,2013,2015,2017. This structure mean indicator in formula

$$w_{ij} = \frac{x_{ij}}{\overline{x}_j}.$$

The presentation in Figure 7 shows that in some countries of the new EU from 2008, the situation is improving (see e.g. Poland marked with a triangle

marker), and in others (see e.g. Latvia and Estonia marked with black markers) situation in relation to 2008 it was worse until 2013. Of course, in 2017, it is better than in 2008, except two countries.

Using the definition of $\|x\|_\mathbf{S}$, different measures of the distance (impossibility) of two vectors can be constructed differently. On the other hand, having the distance measure, you can easily construct both concentration indicators and measurements that order vectors using distances from patterns.

6   Conclusions

This paper discussed techniques for constructing two synthetic indicators (measures) using graphical presentation in the form of radar maps. The work presented for the first time the formal construction of an analogon of a scalar product of vectors used to build indicators and study their properties. The discussed techniques are a typical example of using geometric methods to build synthetic indicators. The work discussed the structures and properties of two indicators: one for ordering objects, another for measuring concentration. Using the definition of $\|\mathbf{x}\|_S$, one can construct different measures of distance (similarity) of two vectors differently. Then, having a measurement of distance, it is easy to construct both concentration indicators and indicators that organize vectors using their distances from templates/models (see in [2]). Due to classic indicators' limited sensitivity to changes (see e.g. Example 4), it is worth including more than one indicator to track changes for the purposes of reporting processes. The presented technique of constructing indicators illustrates how one can construct many different indicators based on other graphic representations and organize a formal study of their properties.

References

[1] R.A. Barnett, M.R. Ziegler, K.E. Byleen and Ch.J. Stocker. *College Mathematics for Business, Economics, Life Sciences, and Social Sciences, 10th ed.* Pearson Prentice Hall, 2005.

[2] Z. Binderman, B. Borkowski, R. Kozera, A. Prokopenya and W. Szczesny. On mathematical modelling of synthetic measures. *Mathematical Modeling and Analysis*, **23**(4):699–711, 2018. https://doi.org/10.3846/mma.2018.042.

[3] Z. Binderman, B. Borkowski and W. Szczesny. Radar coefficient of concentration. *Quantitative methods in economics*, **13**(2):7–21, 2012.

[4] Z. Binderman, W. Szczesny and A. Prokopenya. Radar coefficients of concentrations verifications of properties. *Computer Algebra Systems in Teaching and Research, Siedlce, ed. Collegium Mazovia*, pp. 16–28, 2013.

[5] I.M. Gelfand. *Lectures on linear algebra*. New York, Interscience, 1961.

[6] C. Gini. Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti. A.A.*, **73**(2):1203–48, 1914.

[7] G.J. Glasser. Variance formulas for the mean difference and coefficient of concentration. *Journal of the American Statistical Association*, **57**(229):648–654, 1962. https://doi.org/10.1080/01621459.1962.10500553.

[8] L.D. Hoffmann and G.L. Bradley. *Calculus for Business, Economics, and the Social and Life Sciences, 9th ed.* McGraw Hill, New York, 2007.

[9] T. Kowalczyk, E. Pleszczyńska and F. Ruland. *Grade models and methods for data analysis. Studies in Fuzziness and Soft Computing. Vol. 151.* Springer, Berlin, Heidenberg, New York, 2004. https://doi.org/10.1007/978-3-540-39928-5.

[10] W. Szczesny. Grade correspondence analysis applied to contingency tables and questionnaire data. *Intelligent Data Analysis*, **6**(1):17–51, 2002. https://doi.org/10.3233/IDA-2002-6103.

[11] N.R. Tague. *The quality toolbox, second edition.* ASQ Quality Press, 2005.

[12] A. Wiliński and S. Osowski. Ensemble of data mining methods for gene ranking. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, **60**(3):461–471, 2012. https://doi.org/10.2478/v10175-012-0058-x.

[13] A. Wilinski. Extraction of essential features by quantum density. In R.S. Romaniuk(Ed.), *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2016*, volume 10031, pp. 1528–1534. International Society for Optics and Photonics, SPIE, 2016. https://doi.org/10.1117/12.2249406.

[14] J.Z. Zawistowski, W. Szczesny, B. Borkowski, R. Kozera and Y. Shachmurove. Alternative method of measuring concentration. *Applied Mathematics & Information Sciences*, **10**(1):11–19, 2016. https://doi.org/10.18576/amis/100102.