ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 206, 2007

*Daniel Kosiorowski**

ABOUT PHASE TRANSITIONS IN KENDALL'S SHAPE SPACE

Abstract. In our article, we discuss the choice of space shape, appropriate for describing an economic process and analyze usefulness of the metrics proposed by I. L. Dryden and K. V. Merida (1998). We introduce and interpret the notion of an average shape and its variation for an economic object. We point special attention to the possibility of employing classic tests: T^2 Hotelling of equality of the expected values, of multi-variable analysis of the variance. We compare the proposed approach with a pair of thin plate splines deformation. Theoretical considerations are illustrated with examples of multidimensional economic series.

Key words: Kendall's space of shape, Procrustes analysis, a pair of thin – plate splines.

1. INTRODUCTION – PRELIMINARY DISCUSSION

Ever though intuitive understanding of the shape of a certain object is imprecise, one often uses it when making decisions in his or her everyday life. The shape of a car bodywork signals its owner's prestige, changes in the shape of a child's face signal adulthood, changes in the shape of a ventricle indicate a certain cardiac defect. The problem of changes in the shape of an object with time, being an effect of a certain economic (such as e.g. the affluence of a car owner) or physical factor (e.g. blood pressure, growing older) also has its place in everyday considerations.

The statistical theory of shape (STS) that is being developed nowadays[1] specifies the intuitive comprehension of the shape of an object and makes available a formal apparatus that enables research on reality understood as a realization of a certain multidimensional stochastic model. Within the

* Ph.D., Department of Statistics, Cracow University of Economics, ul. Rakowicka 27, 31-510 Kraków.

[1] The historical background is presented by, among others I. L. Dryden, K. V. Merida, (1998).

theory, the shape of an object belonging to a certain class of objects is defined as the whole information remaining when the location of the object, its scale and rotational effects are removed.

In this work we concentrate on the problem of "phase transition" between shapes. Transitions, we should add, between shapes treated as realizations of a certain stochastic model. Intuitively such transition can be understood as a "significant change of average characteristics" of an object e.g. a face, the mutual position of the numerical characteristics of the shares of a certain stock index, districts of a certain region – under the influence of a change in the intensity of a certain factor. Such research seems to be useful for practical reasons, i.e. for instance when we try to indicate the optimum model of packing for a certain commodity or the arrangement of a store space, or to diagnose cardiac defects more accurately. From the theoretical point of view, STS methods seem to be adequate for an indirect verification of new analytical concepts emerging in the theory of economics, e.g. T. Klecha's concept according to which capital stored in a certain object is described by the ability of this object to perturb a certain space of values, the flow of capital is connected with internal stresses in the substance of the capital carrier.

In this work we present selected[2] analytic STS tools and we demonstrate their usefulness by three empirical examples referring to the relation between the rate of unemployment and the relative wages in districts of Poland, the situation on the stock market, and social opinions about institutions of public life found by an opinion poll.

2. KENDALL'S SHAPE SPACE

Within STS, the shape of an object belonging to a certain class of objects is considered based on the concept of a "marker" – a point which is characteristic of all the objects of the class under consideration and which corresponds with a certain specific substantive (e.g. "fingertips") or mathematical (e.g. "point of high curvature") properties of objects. Usually we consider objects by means of k defined points in the Euclidean space $\mathbf{R}^m$, $k \geqslant 2$. We move each object in such way that its centroid is the origin of coordinates, we normalize their size in such way that the sum of the squares of the distances between the points and the origin was equal to unity.

[2] A rich overview of tools with references to original works can be found for instance in I. L. Dryden, K. V. Merida, (1998); an approach that is different from the one presented here and especially elegant can be found in a work C. R. Rao, S. Suryawanshi (1996), the starting point there is a matrix of Euclidean distances between pairs of markers, the approach is in a sense similar to the classic multidimensional scaling.

As a result of these operations we obtain so called "preshapes" of objects. Any two objects represented by a configuration of k labeled points have the same "shape" if their preshapes can be transformed into one another by means of rotation around a common centroid. As a result of such aliasing of preshapes we receive a set of all possible shapes – "a shape space" denoted by Σ_m^k. "Shape" can be defined as "preshape modulo rotations".

"Kendall's shape spaces" contain the shapes of all possible configurations except those, wherein all points superimpose. We construct them as follows.

We examine an object of a certain class of objects that we are interested in by means of k markers in $\mathbf{R}^m$, $k \geqslant 2$ having the coordinates $x_1^*, \ldots, x_k^*$, which we arrange in a so called $k \times m$ "configuration matrix" X – its rows are the coordinates of the markers. Then we transform orthogonally the k-th row of the configuration matrix as follows:

$$x_0 = \sqrt{k} x_c^* = 1/\sqrt{k}(x_1^* + x_2^* + \ldots + x_k^*),$$

$$\tilde{x}_j = 1/\sqrt{j+j^2}[jx_j^* - (x_1^* + x_2^* + \ldots + x_j^*)] \quad \text{for} \quad 1 \leqslant j \leqslant k-1 \tag{1}$$

As a result we receive matrix $(\sqrt{k}x_c^* \; \tilde{x}_1 \ldots \tilde{x}_{k-1})^{\mathrm{T}}$ – which is equivalent to "multiplying the configuration matrix on the right by so called Helmert's submatrix" (see an example at the end of the paper) produced from Helmert's matrix by deleting its first row. Then we normalize the configuration with respect to size. As a result of these operations we receive matrix $(x_1 \ldots x_{k-1})^{\mathrm{T}}$, which represents the preshape of the object. The shape in itself is represented by this matrix modulo $SO(m)$ – a special orthogonal group operating on the right.

By identifying the space of $(k-1) \times m$ real matrixes with the Euclidean space of the dimension $(k-1) \times m$ as a result of quantity normalization to unity we can recognize that the "preshape" lies on the "unit sphere of dimension $m(k-1)-1$" in this space. The "shape", on the other hand, is identified with "equivalence class" 'orbit' connected with the operation on the right of the special orthogonal group $SO(m)$. In research, it is enough to examine each class by means of its representative, so called "icon".

3. PROKRUST'S PLANAR ANALYSIS

"Prokrust's analysis", which is frequently used in practice, as a measure of differences between two configurations of k defined points in $\mathbf{R}^m$ uses root of the sum of the squares of Euclid distances between corresponding points with optimum matching of two configurations with respect to translation, rotation and scale change. It is a method in which we use a metric induced to the shape space from normalized configuration space $\mathbf{R}^{km}$. Prokrust's analysis is especially easy in two dimensions, where matching with respect to Euclidean similarity can be expressed as a "problem of complex linear regression". In two dimensions, Prokrust's method has a direct solution, with regards to an average shape in the form of an eigenvector. In higher dimensions one should use an appropriate numerical algorithm.

3.1. Distances in shape space

In case of examination of a shape on a plane, when we have two centered configurations:

$$y = (y_1, ..., y_k)^{\mathrm{T}} \quad \text{and} \quad w = (w_1, ..., w_k)^{\mathrm{T}} \quad \text{both in} \quad \mathbf{C}^k$$
$$\text{and} \quad y^* \mathbf{1}_k = 0 = w^* \mathbf{1}_k,$$

it is convenient to consider the following complex regression equation permitting introduction of Prokrust's distances between shapes:

$$y = (a + ib)\mathbf{1}_k + \beta e^{i\theta} w + \varepsilon \tag{2}$$

where $a + ib$ translation, $\beta > 0$ scale, $0 \leqslant \theta \leqslant 2\pi$ angle of rotation, ε $k \times 1$ complex error vector.

"Full Prokrust's distance" between complex configuration w and y is given as

$$d_F(w, y) = \inf_{\beta, \theta, a, b} \left\| \frac{y}{\|y\|} - \frac{w}{\|w\|} \beta e^{i\theta} - a - ib \right\| = \left(1 - \frac{y^* w w^* y}{w^* w y^* y} \right)^{1/2} \tag{3}$$

If we want to account for the "non-Euclidean character of shape space" in our research, we can use so called "Kendall's *rho* distance" $\rho(y, w)$, which is the distance of the nearest large circle between preshapes $\tilde{y} = y/\|y\|$ and $\tilde{w} = w/\|w\|$ on the preshape sphere.

$$\rho(y, w) = \arccos(|\tilde{y}^* \tilde{w}|) \tag{4}$$

3.2. Average shape estimation

In classic statistical analysis, estimation of an average based on a sample generally does not cause problems. In the case of statistical analysis of shape it is not clear what an average means, different circumstances can require different definitions of an average. "Shape spaces are not linear spaces", most of them are not even manifolds. In order to obtain a definition of the average of random variables or probabilistic measures on general spaces or manifolds without linear structure, we introduce[3] a certain number of nonlinear operators, which we use to substitute normal linear ones, e.g. Frechet's average, Cartan's average. For the purpose of this work the following depiction of the problem is sufficient.

Let us consider a situation, where a random sample of a configuration $w_1, \ldots, w_n$ is available from the point of view of a perturbational model.

$$w_i = \gamma_i \mathbf{1}_k + \beta_i e^{i\theta}(\mu + \varepsilon_i), \quad i = 1, \ldots, n \tag{5}$$

$\gamma_i \in \mathbf{C}$ translation vectors, $\beta_i \in \mathbf{R}_+$ scale parameters, $0 \leqslant \theta \leqslant 2\pi$ angle of rotation, $\varepsilon_i \in \mathbf{C}$ – independent errors of zero average, μ average shape in the population.

We obtain the estimator of the "full Prokrust of average shape" $[\hat{\mu}]$ by minimizing, relative to μ, the sum of the squares of full Prokrust's distances from each w_i to unknown average having a unit quantity

$$|\hat{\mu}] = \arg \inf_{\mu} \sum_{i=1}^{n} d_F^2(w_i, \mu) \tag{6}$$

In case of shapes on a plane it is convenient to use the following result: Let us assume that configurations $w_1, \ldots, w_n$ have been centered in such way that $w_i^* \mathbf{1}_k = 0$, then:

Result: We will find full Prokrust's average $[\hat{\mu}]$ as an "eigenvector" corresponding to the largest eigenvalue of the following complex sum of products:

$$S = \sum_{i=1}^{n} w_i w_i^* / (w_i^* w_i) = \sum_{i=1}^{n} z_i z_i^*, \quad \text{where} \quad z_i = w_i / \|w_i\|,$$

$$i = 1, \ldots, n \text{ are preshapes} \tag{7}$$

This solution is unique up to rotation and it corresponds with the estimator of the maximum likelihood (modal shape) under "Bingham's" complex model.

By estimating average Prokrust's shape we receive so called "Prokrust's coordinates" corresponding to the values which the estimated perturbational model (4) takes on. In the planar case, for $w_1, \ldots, w_n$ Prokrust's coordinates are given

$$w_i^P = w_i^* \hat{\mu} w_i / (w_i^* w_i), \quad i = 1, \ldots, n \tag{8}$$

[3] Details of the problem can be found in the following inspiring work D. G. Kendall, D. Barden, T. K. Carne, H. Le, (1999).

To obtain a general measure of shape variability, it is convenient to use the root of the average square of the distance between each configuration and Prokrust's average $[\hat{\mu}]$. We determine this measure $RMS(d_F)$

$$RMS(d_F) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_F^2(w_i, \hat{\mu})} \tag{9}$$

4. STATISTICAL TOOLS

With certain assumptions STS allows the use of classic methods of multidimensional statistics in research. Such analysis is conducted in a space tangent to the preshape sphere at a point usually corresponding with the average shape from the sample. The tangent space is a linearized version of the preshape space, Euklides distance is a good approximation of Prokrust's distance and Kendall's *rho* in the shape space close to the point of contact. There is a certain type of arbitrariness in selection of tangent coordinates, one can use e.g. Prokrust's residuals. Especially noteworthy are, used[4] within STS, modifications of Hotelling's T^2 test, and Goodal's test which refers to analysis of variance, which are used further on in this work.

5. DEFORMATIONS

A global numeric measure of the distances between configurations frequently provides insufficient information about differences between objects. Especially desirable information is one which applies to the nature of local differences between objects. Within the statistical theory of shape, by "global" differences we understood large-scale trends applicable to "all markers". "Local" differences are of smaller scale, i.e. they concern a certain "improper subset of markers". Global differences can be described as smooth changes within a part of the object components. An interesting method for presenting differences between configurations of objects is the calculation of the transformation of the space in which the first object is located into the space of the second object. Such transformation provides to us information about local and global differences in shape. Among many possible forms of the transformation it is worth paying closer attention to so called PTPS[5] (a Pair of Thin-Plane Splines) transformation.

[4] Details of the problem can be found in the following inspiring work I. L. Dryden, K. V. Merida, (1998).

[5] Details of the problem can be found in the following inspiring work I. L. Dryden, K. V. Merida, (1998).

For two $k \times m$ configuration matrixes in $T = (t_1, ..., t_k)^T$ and $Y = (y_1, ... y_k)^T$, **deformation** is a certain bijection of the whole space $\mathbf{R}^m$ w $\mathbf{R}^m$

$$\Phi : \mathbf{R}^m \in t \to (\Phi_1(t), \Phi_2(t), ..., \Phi_m(t))^T \in \mathbf{R}^m \tag{10}$$

which is continuous, smooth, does not allow foliation and for which $y_j = \Phi(t_j)$.

The PTPS deformation, introduced to planar configurations by "Bookstein" is given by the function of two variables

$$\Phi(t) = (\Phi_1(t), \Phi_2(t))^T = c + At + W^T s(t) \tag{11}$$

where t is vector 2×1, $s(t) = (\partial(t - t_1), ..., \sigma(t - t_k))^T$, $k \times 1$

$$\sigma(h) = \begin{cases} \|h\|^2 \log(\|h\|), & \|h\| > 0, \\ 0, & \|h\| = 0. \end{cases}$$

The deformation has $2k + 6$ parameters, i.e.: $c(2 \times 1)$; $A(2 \times 2)$; $W(k \times 2)$. Additionally, $2k$ interpolation limitations: $(y_j)_r = \Phi(t_j)$; $r = 1, 2$; $j = 1, ..., k$, and 6 form limitations: $\mathbf{1}_k^T = 0$; $T^T W = 0$ are introduced.

6. EMPIRICAL EXAMPLES

A. Based on "Labor Statistical Yearbook" 2003 of GUS the following research has been conducted: for year 2001 and for each of Polish provinces 7 districts were drawn and examined considering unemployment rate as recorded on December 31 and the average relative gross wages (Poland = 100). The research was repeated for year 2002, for the same districts that had been drawn earlier. The "class of objects" under consideration "is provinces, markers are districts" in a space of "average unemployment and relative gross wages".

B. The 8 largest companies from WIG20 index were put to examination considering the weekly average percentage increase in their share prices and considering the standard deviation of the percentage increase in share prices in two periods February 16, 2004–September 17, 2004 and July 5, 2003–January 5, 2004. The considered class of objects is "weekly situation on the stock market", "markers are the weekly price increases and the standard deviations" of 8 stocks included in WIG20 index.

C. For years 2000 and 2001 five results each of a poll survey of the assessment of 7 institutions of public life in Poland commissioned by a daily

"Rzeczpospolita". The considered class of objects is "two-months' social assessment of public life institutions", the markers are "percentages of respondents favorably or unfavorably assessing" the examined institutions.

Table 1 presents the quantities of the configuration of the 7 drawn districts of a province and the distance from the province shape to the average shape, additionally it contains the values of the global measure of variability of the province shape calculated using Prokrust's distances and Kendall's *rho* distances.

Table 1

Numerical characteristics of configurations – districts

District	2001		2002	
	Size	*rho*	Size	*rho*
Dolnośląskie	135.79	0.25	137.54	0.27
Kujawsko-Pomorskie	114.53	0.19	118.28	0.27
Lubelskie	96.86	0.22	93.98	0.25
Lubuskie	113.86	0.14	115.28	0.15
Łódzkie	111.28	0.19	107.89	0.23
Małopolskie	126.82	0.17	126.90	0.14
Mazowieckie	109.97	0.26	112.14	0.26
Opolskie	130.96	0.34	130.50	0.37
Podkarpackie	111.11	0.17	112.09	0.18
Podlaskie	110.12	0.15	108.46	0.17
Pomorskie	140.35	0.22	137.38	0.21
Śląskie	97.29	0.23	98.11	0.26
Świętokrzyskie	121.75	0.25	125.64	0.24
Warmińsko-Mazurskie	113.77	0.24	133.38	0.83
Wielkopolskie	120.01	0.22	120.27	0.31
Zachodniopomorskie	113.70	0.17	104.60	0.20
General measure of shape variability	$RMS_p = 0.212$ $RMS_{d_p} = 0.218$			$RMS_p = 0.311$ $RMS_{d_p} = 0.294$

Source: own calculations, GUS's data.

Table 2

Coordinates of average shape – districts 2001

–0.309	–0.021
0.434	–0.021
–0.372	0.040
0.454	0.004
–0.309	–0.001
0.418	–0.006
–0.316	0.006

Source: see Tab. 1.

Table 3

Coordinates of average shape – districts 2002

–0.319	–0.020
0.443	–0.020
–0.362	0.048
0.447	0.030
–0.304	–0.016
0.414	0.005
–0.320	–0.027

Source: see Tab. 1.

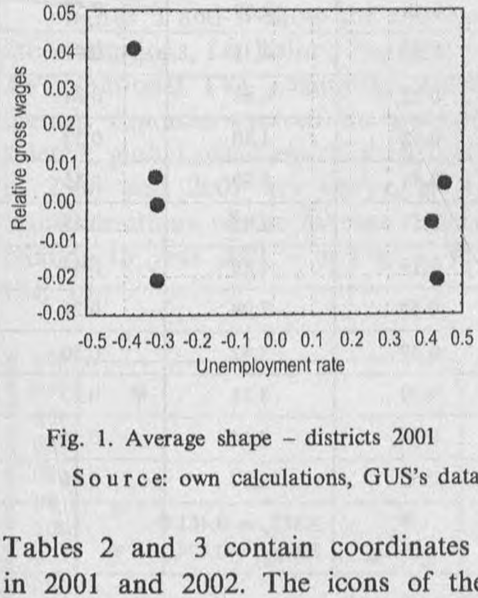


Fig. 1. Average shape – districts 2001

Source: own calculations, GUS's data.

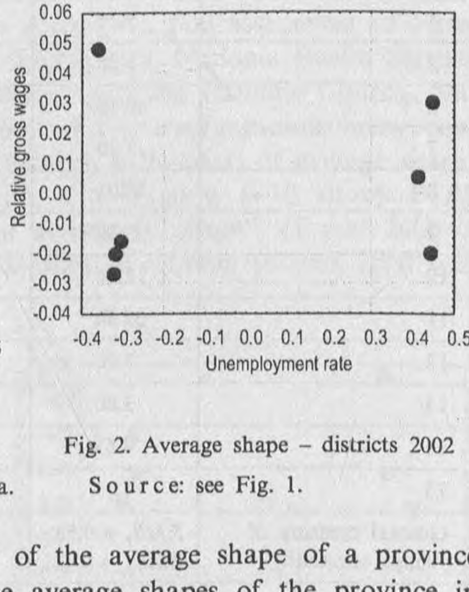


Fig. 2. Average shape – districts 2002

Source: see Fig. 1.

Tables 2 and 3 contain coordinates of the average shape of a province in 2001 and 2002. The icons of the average shapes of the province in 2001 and 2002 are shown in Fig. 1 and 2. Table 4 contains the values of the numerical characteristics of weekly situations on the stock market, considered in the context of the weekly price increase and a weekly standard deviation of the prices of selected stocks included in WIG20 index. Tables 5 and 6 contain the coordinates of average shapes on the stock market in the periods from February 16, 2004 to September 17, 2004 and from July 5, 2003 to January 5, 2004, which is graphically shown in Fig. 3 and 4. Figure 7 presents a PTPS deformation transforming the average shape of a province in 2001 into the average shape of a province in 2002. Figure 8 presents a PTPS deformation converting the average shape of the situation on the stock exchange in February 16, 2004–September 17, 2004 into the average shape of the situation on the stock exchange in July 5, 2003–Janurary 5, 2004.

Table 4

Numerical characteristics of configurations – WIG20

Week	Size	*rho*	Size	*rho*
	04.02.16–04.09.17		05.07.03–05.01.04	
1	2.25	0.69	4.79	0.37
2	2.37	0.57	4.74	0.67
3	2.53	0.35	2.93	0.52
4	3.09	0.48	2.69	0.59
5	3.26	0.46	4.31	0.41
6	60.48	0.12	6.61	0.24
7	3.10	0.42	4.40	0.32
8	3.05	0.48	5.80	0.31
9	3.36	0.67	2.85	0.45
10	13.70	1.19	4.89	0.44
11	65.34	0.19	7.04	0.27
12	2.96	0.39	3.42	0.50
13	3.60	0.59	3.21	0.53
14	4.65	0.35	3.64	0.34
15	4.23	0.50	2.47	0.66
General measure of shape variability	$RMS_\rho = 0.551$ $RMS_{d_F} = 0.461$		$RMS_\rho = 0.461$ $RMS_{d_F} = 0.434$	

Source: own calculations.

Table 5

Average WIG20 configuration: 04.02.16–04.09.17

–0.272	0.004
0.358	0.004
–0.386	0.008
0.342	–0.037
–0.342	0.049
0.279	–0.009
–0.398	–0.007
0.418	–0.011

Source: own calculations.

Table 6

Average WIG20 configuration: 05.07.03–05.01.04

–0.349	0.016
0.294	0.016
–0.390	0.000
0.364	–0.003
–0.364	0.037
0.286	0.073
–0.284	–0.063
0.442	–0.077

Source: own calculations.

Figures 5 and 6 show the average shape of a poll assessment of social life institutions, i.e. Police, President, Government, National Health Service, TVP (national TV), Commune Authorities and the Catholic Church. Numerical characteristics of the quantities of the configurations being considered, global measures of shape variability, estimations of average shapes in 2000 and 2001 are shown in Tab. 7–9. Figures 9–10 shows PTPS transformations of the average "social assessment shape": of year 2000 in relation to year 2001 – in Fig. 9, year 2001 in relation to year 2000 – in Fig. 10.

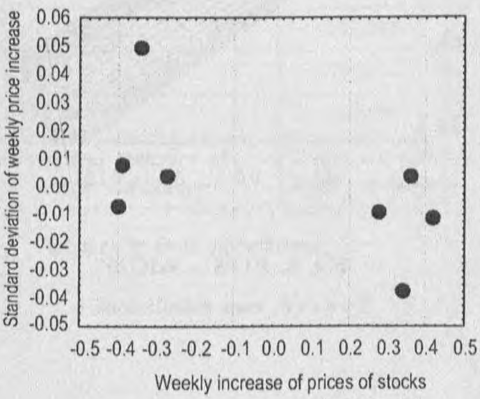


Fig. 3. Average shape of WIG20 – Feb. 16, 2004–Sep. 17, 2004

Source: own calculations.

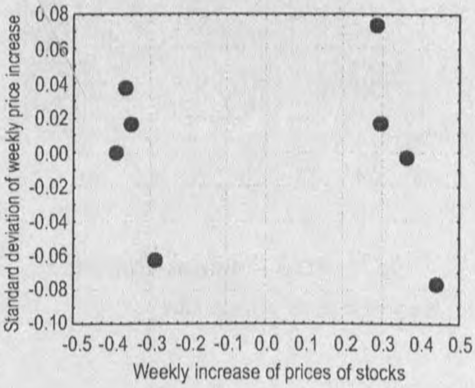


Fig. 4. Average shape of WIG20 – July 5, 2003–Jan. 5, 2004

Source: own calculations.

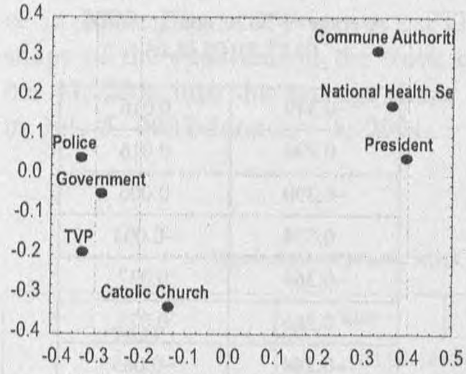


Fig. 5. Average shape – social assessments of public institutions 2000

Source: own calculations.

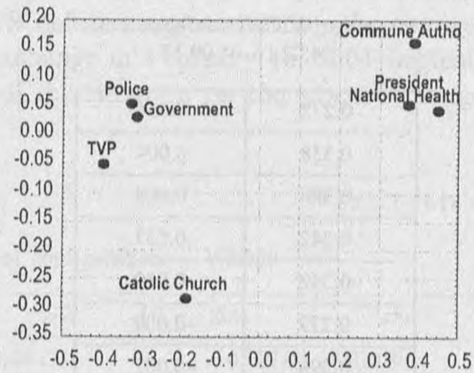


Fig. 6. Average shape – social assessments of public institutions 2001

Source: own calculations.

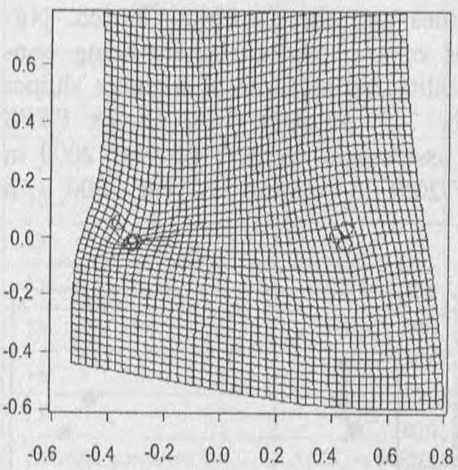


Fig. 7. PTPS – districts 2001/2002

Source: own calculations.

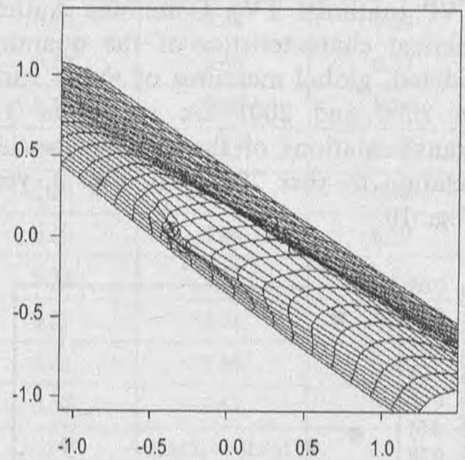


Fig. 8. PTPS – WIG20

Source: own calculations.

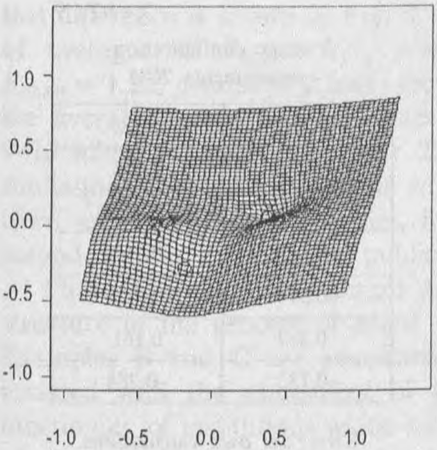


Fig. 9. PTPS – assessments 2000/2001

S o u r c e: own calculations.

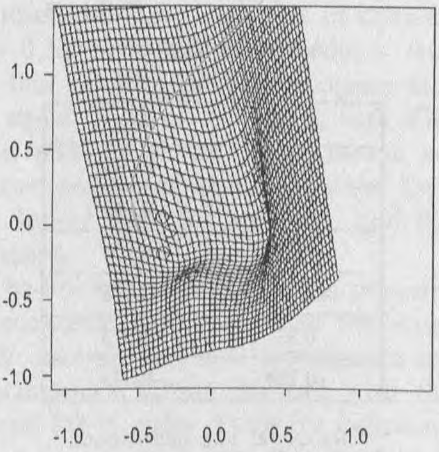


Fig. 10. PTPS – assessments 2001/2002

S o u r c e: own calculations.

T a b l e 7

Numerical characteristics of configurations – social assessments of public institutions

Period	Size	*rho*	Size	*rho*
	2000		2001	
1	55.513	0.692	60.886	0.38
2	88.048	0.506	97.583	0.411
3	107.974	0.503	105.134	0.39
4	82.977	0.335	90.043	0.271
5	63.154	0.32	58.982	0.235
General measure of shape variability	$RMS_\rho = 0.490$ $RMS_{d_p} = 0.466$		$RMS_\rho = 0.345$ $RMS_{d_p} = 0.337$	

S o u r c e: own calculations.

Table 8

Average configuration – assessments 2000

–0.334	0.045
0.396	0.045
–0.285	–0.049
0.36	0.174
–0.33	–0.195
0.33	0.313
–0.138	–0.333

Source: own calculations.

Table 9

Average configuration – assessments 2001

–0.323	0.052
0.375	0.052
–0.311	0.029
0.449	0.043
–0.394	–0.053
0.387	0.161
–0.183	–0.284

Source: own calculations.

7. RESULTS AND SUMMARY

We can interpret the size of the configuration of a province's districts as a characteristic measure of that province's uniformity, the distance between a province and the average shape – as a global measure of distinction, relating to disproportions prevailing in the province[6]. In Fig. 1 and 2 we can see 3 clusters: i) high unemployment and low wages, ii) low unemployment and low wages, iii) one district with high unemployment and high wages. Year 2002 in relation to 2001 indicates, the cluster elements: i) overlapped, the economic situation deteriorated, the cluster elements, ii) grew away from each other, iii) relative wages grew with increase in unemployment. The qualitative character of these changes can be seen in Fig. 7, where a relevant PTPS deformation is presented. The values of global variability measures *RMS* indicate a bigger diversity between provinces in 2002 than in 2001. The results of Hotelling's test of equality of averages $T^2 : F_{10.21} = 0.21$, p-value $= 0.9931$ (Goodall's statistics value: $F_{10.300} = 0.2$, p-value $= 0.996$) which, provided that the test assumptions are fulfilled, means that despite local differences, the globally average shape of a province in 2001 can be considered to be the same as in 2002. From Tab. 4 we may read that the global measure of the weekly variation of the situation on the stock market in the period February 16, 2004–September 17, 2004 was higher than in the period July 5, 2003–January 5, 2004, the periods differ in the "stability of the weekly situation". In Fig. 3 and 4 we can see a difference in the average shape of WIG20, the qualitative character of

[6] The results can be compared with those obtained using classic methods, e.g. based on the work K. Zając, D. Kosiorowski (2004)

this difference is shown in Fig. 8. The results of Hotelling's test of equality of averages $T^2: F_{12.17} = 1.27$, p-value = 0.3153 (results of Goodall's test: $F_{12.336} = 1.22$, p-value = 0.2668) indicate that we may globally recognize that the average shapes of the indexes are equal. Figures 5–6 along with Fig. 9–10 allow to assert that year 2000 in relation to year 2001 means assimilation of social assessments within two groups of institutions, the first: town and commune authorities, health service and the president, and the second: police, government, public television.

Phase transitions for example **A** can be linked to changes in the province structure in the process of social and economic development of the state. Examples **B** and **C** are connected with changes in social atmosphere associated with the assessment of the economy's future (**B**) and with the functioning of institutions which keep social life in order. Tools for indicating phase transitions can be sought either in the two presented statistical texts or in the proposed PTPS transformation. The proposed "tests" are used to select changes of a "global character", the "local" character of changes is shown by "PTPS" deformation.

Despite the fact that no global phase transitions were observed in the examples investigated in this work, PTPS transformations demonstrate how diverse locally can be a global lack of changes. PTPS transformation permits a possibility of modification allowing for emphasizing the variability of a certain especially interesting set of markers, which seems to be an interesting direction for further research. This work is only the first step in the direction of a very interesting, in the author's opinion, area of statistical research concerning relations between global and local properties of economic systems, and extracting characteristics, which so far escaped economic examination, from economic systems.

The statistical theory of shape can be an especially valuable approach both in practice and in theoretical research. A shortcoming of its models can be high restrictions concerning the assumptions for the examined phenomenon. Yet this shortcoming can motivate researchers to propose tolerant non-parametric[7] methods of multidimensional statistics that would be adequate to shape problems. The effort is worthwhile, as the picture of social and economic reality emerging with the application of the statistical theory of shape is not obvious at the moment of bringing up a problem. This determines the pleasure of cognition.

Calculations have been made by means of I. Dryden's "The shapes Package"made available under GNU license on R project pages.

---

[7] A noteworthy direction of research, in the light of the subject matter presented in this work, are the methods presented inter alias in the work D. Kosiorowski (2004a).

REFERENCES

Dryden I. L., Merida K. V. (1998), *Statistical Shape Analysis*, John Wiley & Sons, New York.

Goodall C. R., Marida K. V. (1999), *Projective Shape Analysis*, "Journal of Computational and Graphical Statistics", **8**, 2, 143–168.

Kendall D. G., Barden D., Carne T. K., Le H. (1999), *Shape and Shape Theory*, John Wiley & Sons, New York.

Kosiorowski D. (2004a), *Data Depth Concept in Multivariate Financial Time Series*, 11th International scientific conference "Statistics in management of socio economic development", October 20–22, Kyiv National Economic University.

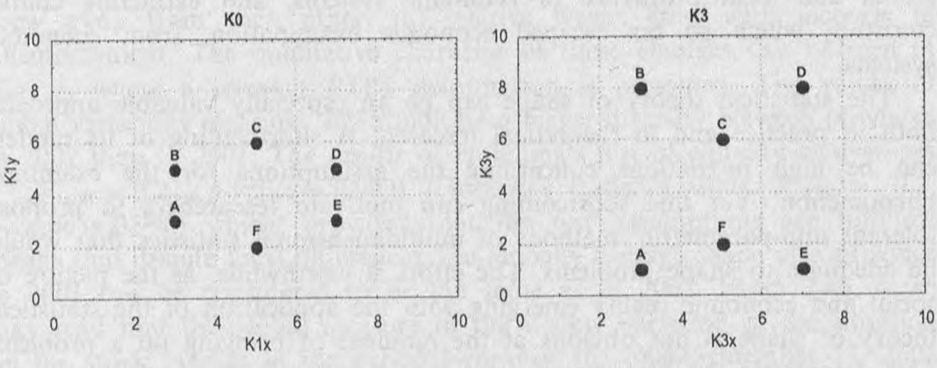
Kosiorowski D. (2004b), *Individual Rationality Versus Group Rationality in Statistical Modelling Issues*, [in:] D. Baier, K.-D. Wernecke (eds.), *Innovations in Classification, Data Science, and Information Systems. Proc. 27th Annual GfKl Conference*, Springer-Verlag, Heidelberg–Berlin, 239–248.

Rao C. R., Suryawanshi S. (1996), *Statistical analysis of object based on landmark data*, Proc. Natl. Acad. Sci. USA, **93**.

Zając K., Kosiorowski D. (2004), *Przemiany kulturowe a rozwój społeczno-gospodarczy*, [in:] *Klasyfikacja i analiza danych – teoria i zastosowania*, K. Jajuga, M. Walesiak (red.), "Taksonomia", **12**, (Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 1076, Wydawnictwo AE we Wrocławiu), 26–37.

8. AN EXAMPLE

Let us consider two configurations $k = 6$ points in $m = 2$ dimensions: **K0** and **K3**



For the above configurations we have two corresponding configuration matrices:

$$\mathbf{K0}^{\mathrm{T}} = \begin{bmatrix} 3 & 3 & 5 & 7 & 7 & 5 \\ 3 & 5 & 6 & 5 & 3 & 2 \end{bmatrix} \qquad \mathbf{K3}^{\mathrm{T}} = \begin{bmatrix} 3 & 3 & 5 & 7 & 7 & 5 \\ 1 & 8 & 6 & 8 & 1 & 2 \end{bmatrix}$$

Or using complex notation we have two configuration vectors:

$$\mathbf{k0} = [3+3i, 3+5i, 5+6i, 7+5i, 7+3i, 5+2i],$$

$$\mathbf{k3} = [3+1i, 3+8i, 5+6i, 7+8i, 7+1i, 5+2i].$$

Centroids of the configurations are given as:

$$\overline{\mathbf{K0}} = [5,4], \quad \overline{\mathbf{K3}} = [5,4,3].$$

Centroid sizes of the configurations are:

$$S(\mathbf{K0}) = 5.291, \quad S(\mathbf{K3}) = 8.563.$$

In order to filter a location of the configurations we multiply them from left by 5×6 Helmert submatrix $\mathbf{H}$:

$$\mathbf{H} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 & 0 & 0 \\ -1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} & 0 & 0 & 0 \\ -1/\sqrt{12} & -1/\sqrt{12} & -1/\sqrt{12} & 3/\sqrt{12} & 0 & 0 \\ -1/\sqrt{20} & -1/\sqrt{20} & -1/\sqrt{20} & -1/\sqrt{20} & 4/\sqrt{20} & 0 \\ -1/\sqrt{30} & -1/\sqrt{30} & -1/\sqrt{30} & -1/\sqrt{30} & -1/\sqrt{30} & 5/\sqrt{30} \end{bmatrix}.$$

We have:

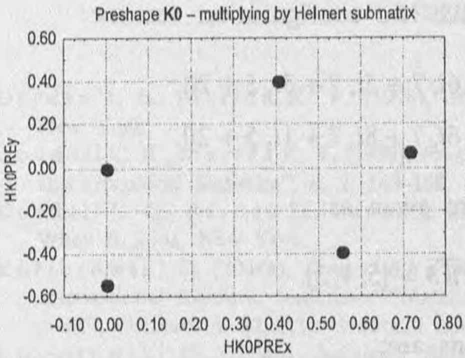
$$(\mathbf{HK0})^{\mathrm{T}} = \begin{bmatrix} 0 & 1.63 & 2.89 & 2.24 & 0 \\ 0 & 1.63 & 0.29 & -1.57 & -2.19 \end{bmatrix},$$

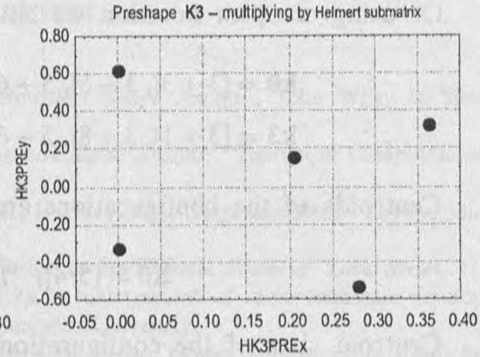
$$(\mathbf{HK3})^{\mathrm{T}} = \begin{bmatrix} 0 & 1.63 & 2.89 & 2.24 & 0 \\ 4.95 & 1.22 & 2.60 & -4.25 & -2.56 \end{bmatrix}$$

and dividing the matrices **HK0** and **HK3** by their centroid sizes (4.026 and 7.98) we obtain preshapes:

$$(\mathbf{HK0}_{\mathrm{pre}})^{\mathrm{T}} = \begin{bmatrix} 0 & 0.4 & 0.72 & 0.56 & 0 \\ 0 & 0.4 & 0.07 & -0.39 & -0.54 \end{bmatrix},$$

$$(\mathbf{HK3}_{\mathrm{pre}})^{\mathrm{T}} = \begin{bmatrix} 0 & 0.21 & 0.36 & 0.28 & 0 \\ 0.62 & 0.15 & 0.33 & -0.53 & -0.32 \end{bmatrix}.$$





In order to filter the location of the configuration we can also use a centering matrix. We multiply the configuration matrices from left by the 6×6 centering matrix $\mathbf{C} = \mathbf{1}_6 - 1/6 \cdot \mathbf{1}_6\mathbf{1}_6^{\mathrm{T}}$.

We have the centered configuration matrices:

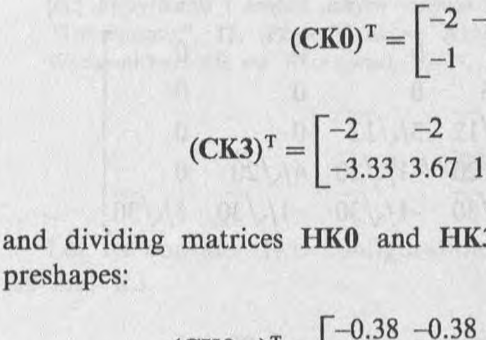
$$(\mathbf{CK0})^{\mathrm{T}} = \begin{bmatrix} -2 & -2 & 0 & 2 & 2 & 0 \\ -1 & 1 & 2 & 1 & -1 & -2 \end{bmatrix},$$

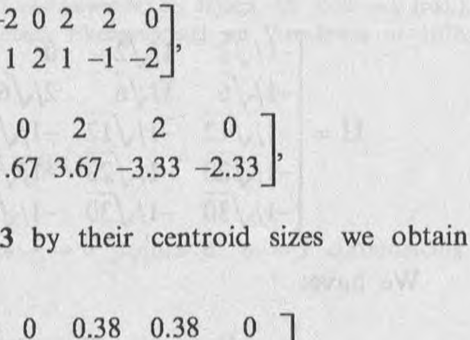
$$(\mathbf{CK3})^{\mathrm{T}} = \begin{bmatrix} -2 & -2 & 0 & 2 & 2 & 0 \\ -3.33 & 3.67 & 1.67 & 3.67 & -3.33 & -2.33 \end{bmatrix},$$

and dividing matrices **HK0** and **HK3** by their centroid sizes we obtain preshapes:

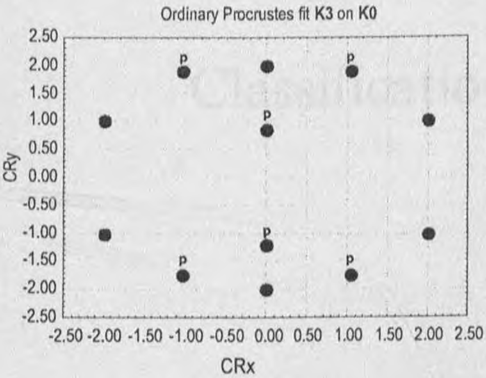
$$(\mathbf{CK0}_{\mathrm{pre}})^{\mathrm{T}} = \begin{bmatrix} -0.38 & -0.38 & 0 & 0.38 & 0.38 & 0 \\ -0.19 & 0.19 & 0.38 & 0.19 & -0.19 & -0.38 \end{bmatrix},$$

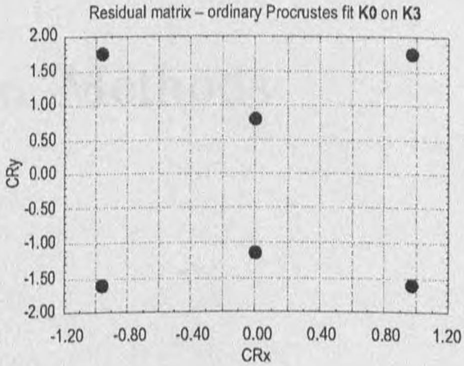
$$(\mathbf{CK3}_{\mathrm{pre}})^{\mathrm{T}} = \begin{bmatrix} -0.23 & -0.23 & 0 & 0.23 & 0.23 & 0 \\ -0.39 & 0.43 & 0.19 & 0.43 & -0.39 & -0.27 \end{bmatrix},$$





An ordinary Procrustes analysis is Least Squares fit one configuration to another subject to Euclidean similarity transformations (scale β, rotation Γ and translation γ). For the configurations **K0** and **K3** we have: estimated rotation martix $\hat{\Gamma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, scale parametr $\hat{\beta} = 0.518$ and translation parametr: $\gamma = (0, 0)$. Ordinary sum of squares equals $OSS = 8.309$ and Kendall's distance between **K0** and **K3** is $\rho(\mathbf{K0}, \mathbf{K3}) = 0.576$. A residual matrix: $\mathbf{R} = \mathbf{K3} - \mathbf{K0}^P$ and Procrustes coordinates (Procrust fit) $\mathbf{K0}^P$ are given below.





$$(\mathbf{K0}^P)^T = \begin{bmatrix} -1.036 & -1.036 & 0 & 1.036 & 1.036 & 0 \\ -1.727 & 1.9 & 0.863 & 1.9 & -1.727 & -1.209 \end{bmatrix},$$

$$(\mathbf{R})^T = \begin{bmatrix} -0.964 & -0.964 & 0 & 0.964 & 0.964 & 0 \\ -1.606 & 1.767 & 0.804 & 1.767 & -1.605 & -1.124 \end{bmatrix}.$$

Daniel Kosiorowski

O PRZEJŚCIACH FAZOWYCH W PRZESTRZENI KSZTAŁTU KENDALLA

W pracy dyskutujemy zarówno na temat wyboru właściwej przestrzeni kształtu dla opisu procesu ekonomicznego, jak i użyteczności metryk proponowanych przez I. L. Drydena i K. V. Meridę (1988). Wprowadzamy i interpretujemy pojęcie przeciętnego kształtu oraz wariancji kształtu ekonomicznego. Zwracamy szczególną uwagę na możliwość wykorzystania klasycznych testów: równości wartości oczekiwanych T^2 Hotellinga i wielozmiennej analizy wariancji. Rozważania teoretyczne ilustrujemy na przykładach wielowymiarowych szeregów finansowych.