On Structural Analysis of Parliamentarian Voting Data

Tomas KRILAVIČIUS

Informatics Faculty, Vytautas Magnus University
Vileikos 8, Kaunas, Lithuania
e-mail: t.krilavicius@gmail.com

Antanas ŽILINSKAS

Institute of Mathematics and Informatics
Akademijos st. 4, Vilnius, Lithuania
e-mail: antanasz@ktl.mii.lt

Abstract. We investigate applicability of quantitative methods to discover the most fundamental structural properties of the most reliable political data in Lithuania. Namely, we analyze voting data of the Lithuanian Parliament. Two most widely used techniques of structural data analysis (clustering and multidimensional scaling) are compared. We draw some technical conclusions which can serve as recommendations in more purposeful application of these methods.

Keywords: clustering, multidimensional scaling.

1. Introduction

Quantitative methods constitute a well established tool box of Western political researchers. Besides of academic sphere quantitative methods are used also, e.g., by organizations that monitor parliaments and other governmental institutions (Irving and Todd, 2007). Contrary, in Eastern Europe quantitative analysis of political behavior is a rather new research area. Possibly, analysts tend to avoid quantitative methods because of general fuzziness in application of traditional political concepts in new democratic states. However, the contrary reaction would be more constructive: wider application of quantitative methods would extend description of actual political behavior by correct theoretical terms.

To investigate some technical aspects of applicability of quantitative methods to the investigation of behavior of Lithuanian political parties we consider the most fundamental structural properties of the most reliable data. Namely, we analyze voting data of the Lithuanian Parliament. Two most widely used techniques of structural data analysis (clustering and multidimensional scaling) are compared. We draw some technical conclusions which can serve as recommendations in more purposeful application of these methods.

Table 1

Lithuanian parliamentarian parties

Abbr.	Party name	Party name in Lithuanian
DP	Labour party	Darbo partija
LSDP	Lithuanian Social Democratic Party	Lietuvos socialdemokratų partija
NS	New Union (Social Liberals)	Naujoji sąjunga (socialliberalai)
VLP	Peasants and People's	Valstiečių liaudininkų partija
PDP	Civil Democracy (faction)	Pilietinės demokratijos grupė
MG	Mixed Group (not party)	Mišri Seimo narių grupė
TTLD	Order and Justice (Liberal Democrats)	Tvarka ir Teisingumas (liberalai demokratai)
LCS	Liberal and Center Union	Liberalų ir centro sąjunga
LP	Liberals Movement	Liberalų sąjūdis
TS	Homeland Union	Tėvynės sąjunga

2. Data to Analyze

The Lithuanian Parliament, Seimas, after election of 2004 is constituted of 10 factions under umbrellas of corresponding parties; they are listed in Table 1. Parties have declared their political, economic, etc. goals in their programmes. The actual political behavior, however, not always corresponds to the declarations. This research aims to highlight proximity/dissimilarity of parties based not on declarations but on the most important aspect of political behavior, i.e., parliamentarian voting. The discovered structure is compared to the theoretically expected one. Such a structural analysis is interesting to us from the point of view of possible extension of application of the considered methods to the relatively new area of political data of new democratic states; the extension can be challenging, and we will discuss needed enhancing of the known methods.

For this investigation data has been collected from the Lithuanian Parliament web page (Seimas2007):

- only roll call voting[1] is considered,
- voting information over period from 2005-01-12 to 2007-01-16 is used (term of office is Seimas 2004–2008),
- at overall 243 plenary sittings (numbers 1379–1621).

The number of members of parliament (MP) is equal to 141; however the data of voting of 138 members, whose membership covered all the considered period, is taken into account. Below MPs are coded by numbers; for decoding (name, party membership) we refer to (Krilavičius *et al.*, 2007).

Only voting with more than 20% disagreement was selected for analysis to emphasize differences among MPs (however, some experiments were performed with whole data set, but we do not discuss them in this paper). As a consequence of such choice, 370 roll calls out of 3838 were taken into account, while others were ignored.

---

[1] A *roll call vote* is a vote held on the record. The name of the Representative and her/his voting position are noted together.

We specify the following possible voting behavior of MP's:

1) `yes`;
2) `no`;
3) `abstain`;
4) did not participate in the plenary sitting (during voting);
5) did not vote, but participated in the plenary during voting.

It is easy to see that voting is not just a binary procedure. Even standard MP voting behavior amounts to the three kinds, i.e., `yes`, `no` and `abstain`. Furthermore, no so rarely MPs exhibit the following behavior: they do not participate in the plenary sitting during the voting or they do not vote at all.

There is no standard technique for mapping such a data to numerical values, e.g., (Hix *et al.*, 2006) use several different mappings for Europarliament voting analysis. In (Irving and Todd, 2007) slightly different mapping is used where voting together with party and against party has different weights:

1) `yes` (aye) maps to 2;
2) `yes` while voting with MP's party (tellaye) maps to 1;
3) `no` (no) maps to 4;
4) `no` while voting with MP's party (tellno) maps to 5;
5) `abstain` (both) maps to 3;
6) Non-participation maps to -9.

Notions in the parentheses (aye, tellaye, no, tellno, both) are the original notions used in the (Irving and Todd, 2007).

Of course, other mappings can be defined. Political analysts can choose appropriate mappings based on their intuition, theoretical considerations and results of preliminary experimentation. For this technical investigation we have chosen rather simple mapping reflecting that `yes` and `no` are opposite while other votes are somewhere in between, e.g., `abstain` is exactly between the `yes` and `no`:

1) `yes`$= 1$;
2) `no`$= -1$;
3) `abstain`$= 0$;
4) Non-voting equals to 0.

The voting matrix $V(\cdot, \cdot)$ corresponding to the discussed data and notation is presented in (Krilavičius *et al.*, 2007), where $V(k, i)$ denotes the voting result of kth MP at ith roll call.

2.1. *Natural Embedding of Data*

We are going to highlight the structure of Lithuanian parliament defined by voting data and to compare it to the formal structure defined by the membership of MP to parties. In both cases it is a structure of a set of N object, where $N = 138$ is the number of considered MPs. For the analysis of a structure of a set of objects a proximity/dissimilarity measure of the objects is needed.

The formal party based structure of the parliament is defined by a simplest dissimilarity measure of MPs: dissimilarity between members of the same parliamentary faction is equal to 0, and dissimilarity between members of different factions is equal to 1. Mathematically the simplest dissimilarity measure on a set of objects can be interpreted as an equivalence relation implying factorization of the set, and factor sets represent a structure of the considered set. Therefore natural embedding of minimal but a structure implying data about MPs is in a finite set factorized according to the equivalence relation where factor sets correspond to parliamentary factions.

However, actual political behavior not always is fully defined by the party membership. The most important official activity of a MP is voting. Therefore embedding of the most rich data about a MP is in s-dimensional vector space where $s = 370$ is number of the considered roll calls. To discover the voting based structure of the parliament we consider MPs as points in s-dimensional vector space, and dissimilarity between MPs is defined by means of a metric over this space. Usually, Minkowski metric is used, and dissimilarity between kth and mth MP is evaluated as Minkowski (p-norm) distance

$$d_p(k, m) = \left(\sum_{i=1}^{s} \big| V(k, i) - V(m, i) \big|^p \right)^{1/p}, \tag{1}$$

where $V(\cdot, \cdot)$ is the voting matrix, and $p \geqslant 1$ is a parameter.

In this paper we use dissimilarities defined by two most frequently used distances:

- 1-norm distance, also known as *Manhattan* or *city block* distance

$$d_1(k, m) = \sum_{i=1}^{s} \big| V(k, i) - V(m, i) \big|, \tag{2}$$

- 2-norm distance widely known as *Euclidean distance*

$$d_2(k, m) = \sqrt{\sum_{i=1}^{s} \big(V(k, i) - V(m, i) \big)^2}. \tag{3}$$

For the discussion on choice of a metric in the problems of visualization of multidimensional data we refer to (Žilinskas and Žilinskas, 2006).

3. Data Reduction for Structural Analysis

Below we investigate applicability to our problem of two popular methods of structural analysis: cluster analysis (Theodoridis and Koutroumbas, 2006) and multidimensional scaling (Borg and Groenen, 2005; Cox and Cox, 2001). In both cases a set of objects characterized by pairwise dissimilarities is considered. Particularly the objects can be points in a multidimensional vector space where dissimilarity is measured using a metric defined in this space. For the further analysis the data presented as voting matrix should

be reduced to a dissimilarity matrix $D_p = (d_p(k, m))$ where dissimilarities $d_p(\cdot, \cdot)$ can be measured by both, city block ($p = 1$) and Euclidean ($p = 2$) distances.

The reduction of data is illustrated by the following example. Let there are three MPs and they have voted four times. Let the first MP all the time voted `yes`, the second MP `yes`, `no`, `yes` and `no`, and the third has chosen to `abstain` the first two times, voted `yes` third time, and has skipped the last voting. The voting results are represented by the following voting matrix

$$V = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4}$$

The date presented in matrix (4) applying formulae (2) and (3) is reduced to the following dissimilarity matrices

$$D_1 = \begin{pmatrix} 0 & 4 & 3 \\ 4 & 0 & 3 \\ 3 & 3 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 0 & 2.83 & 1.73 \\ 2.83 & 0 & 1.73 \\ 1.73 & 1.73 & 0 \end{pmatrix}. \tag{5}$$

4. Results of Cluster Analysis

Clustering is a technique for indicating not intersecting subsets (called clusters) of a given set of objects in such a way that the objects in the same subset are more similar than the objects belonging to the different subsets (Theodoridis and Koutroumbas, 2006; pp. 541–587). The prerequisite for clustering is a measure of similarity/dissimilarity defined over pairs of objects. However this measure does not define a unique partitioning of the considered set of objects. Generally speaking, results of clustering are method dependent, particularly they are dependent on some heuristically chosen parameters.

Hierarchical clustering algorithms produce a hierarchy of nested clusters (Theodoridis and Koutroumbas, 2006; p. 541–587).

Basically, there are two categories of *hierarchical clustering* algorithms: *agglomerative* and *divisive*. The first starts from the single objects as clusters and groups them by merging two most similar (by a predefined criteria) at a step, and ultimately ends-up with a hierarchy of clusters closed by one cluster containing all objects. The last one follows the inverse path, i.e., by splitting initially one cluster of objects to a set of clusters consisting of one object, in such a way producing a whole hierarchy of clusters.

In this paper we apply agglomerative clustering, which we present concisely here. Let $x_i \in \mathbb{R}^s$, $i = 1 \ldots N$ be a set of objects (points in an s-dimensional space). Dissimilarity between x_i, x_j is measured by means of Minkowski distance $d_p(x_i, x_j)$. Let $\mathcal{C}_k$ and $\mathcal{C}_l$ be two clusters, i.e., two not intersecting subsets of the set of x_i, $i = 1 \ldots N$. We will measure the distance between clusters by means of *furthest neighbor*

$$\Delta_p(\mathcal{C}_k, \mathcal{C}_l) = \max_{x_i \in \mathcal{C}_k, x_j \in \mathcal{C}_l} \{d_p(x_i, x_j)\}. \tag{6}$$

An abstract algorithm (following lines of *Generalized Agglomerative Scheme* (Theodoridis and Koutroumbas, 2006)) can be defined as follows.

- Choose an initial clustering $\mathfrak{R}_0 = \{\mathcal{C}_i = \{x_i\}, i = 1 \ldots N\}$ and set a step $t = 0$.
- Repeat until all objects lie in a single cluster.
  - Increment step $t = t + 1$.
  - Among all possible pairs of clusters find the pair of minimal distance $(\mathcal{C}_i, \mathcal{C}_j)$.
  - Merge the selected pair resulting in the new cluster $\mathcal{C}_q = \mathcal{C}_i \cup \mathcal{C}_j$. Produce new clustering $\mathfrak{R}_t = (\mathfrak{R}_{t-1} \setminus \{\mathcal{C}_i, \mathcal{C}_j\}) \cup \{\mathcal{C}_q\}$ by removing clusters $\mathcal{C}_i, \mathcal{C}_j$ from the clustering and adding new cluster $\mathcal{C}_q$.

The outcome of experiments depends on the proximity measure used to calculate distances between clusters, e.g., *single linkage* or *complete linkage*. In this investigation only *complete linkage*, corresponding to (6), is used since in the case of small distances single linkage often implies *chaining phenomena*, i.e., objects are added to the same cluster one by one until all of them lie in the same single cluster. Early experiments with voting data produced the above mentioned effect, and therefore the complete linkage was chosen.

Clustering was performed using both, city block ($d_1(\cdot, \cdot)$) and Euclidean ($d_2(\cdot, \cdot)$) dissimilarities in voting space. Implementation of clustering and distance functions from Matlab Statistics Toolbox 6 (StatToolbox2007) was used. Experiments with full and selected set of roll calls were performed, using both single and complete linkage proximity measures, but because of space limitation we present only the most interesting results.

The results of clustering (2 and 4 clusters, city block and Euclidean dissimilarity measures) are presented in Figs. 1 and 2, and Table 2. *Dendrograms* (Figs. 1 and 2) illustrate hierarchical arrangement of clusters produced by clustering algorithm. Leaves represent data points or clusters, and upsidedown U-shaped lines connect the leaves forming new clusters. The height of the lines represents the distance between the objects being connected. Here dendrograms illustrate the order of clusters merging, i.e., in both, Euclidean and city block, cases, the first and the second clusters are merged, then the third and fourth follow, see Table 2 for clusters.
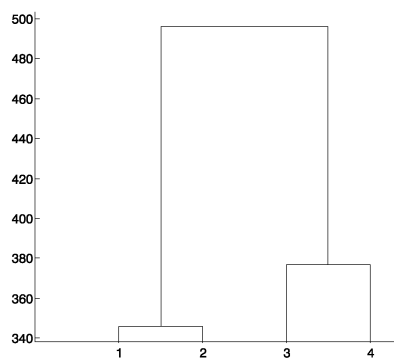

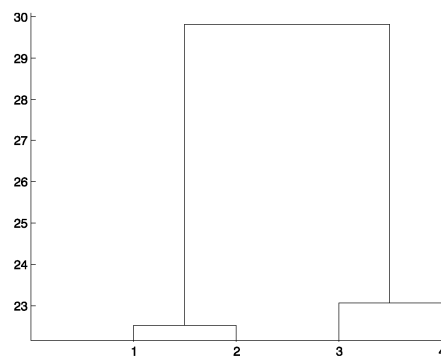Fig. 1. Dendrogram for clustering with city block distances.

Fig. 2. Dendrogram for clustering with Euclidean distances.

Table 2

Clustering of voting results, 4 clusters

City block	Euclidean
Cluster: 1; Cluster size: 9; **MG**(1): 9; **PDF**(1): 67; **TTLDF**(7): 1 33 56 68 86 129 134;	Cluster: 1; Cluster size: 10; **LSDPF**(1): 108; **MG**(1): 40; **PDF**(1): 67; **TTLDF**(7): 1 33 56 68 86 129 134;
Cluster: 2; Cluster size: 45; **LCSF**(8): 13 21 22 60 61 81 99 123; **LF**(11): 3 14 23 32 38 41 54 65 118 121 124; **MG**(1): 144; **TSF**(25): 2 4 10 24 25 27 28 31 34 45 46 48 58 59 63 66 88 96 98 104 115 116 130 138 139;	Cluster: 2; Cluster size: 45; **LCSF**(8): 13 21 22 60 61 81 99 123; **LF**(11): 3 14 23 32 38 41 54 65 118 121 124; **MG**(1): 144; **TSF**(25): 2 4 10 24 25 27 28 31 34 45 46 48 58 59 63 66 88 96 98 104 115 116 130 138 139;
Cluster: 3; Cluster size: 27; **LF**(1): 141; **LSDPF**(14): 5 11 19 50 53 55 70 77 82 103 106 108 133 143; **NSF**(1): 128; **PDF**(3): 78 97 146; **VLF**(8): 9 73 76 93 95 101 114 125;	Cluster: 3; Cluster size: 46; **LF**(1): 141; **LSDPF**(22): 5 11 16 19 49 50 53 55 70 77 82 83 90 100 103 106 107 110 120 133 142 143; **MG**(1): 112; **NSF**(9): 44 52 74 84 105 109 113 117 128; **PDF**(3): 78 97 146; **VLF**(10): 9 73 76 89 93 95 101 111 114 125;
Cluster: 4; Cluster size: 57; **DPF**(24): 8 15 17 18 26 30 35 37 39 51 57 64 71 72 75 91 92 119 122 135 136 137 140 145 147; **LS-DPF**(12): 16 43 49 83 85 90 100 107 110 120 132 142; **MG**(1): 112; **NSF**(8): 44 52 74 84 105 109 113 117; **PDF**(7): 7 42 47 62 126 131 148; **VLF**(4): 29 79 89 111;	Cluster: 4; Cluster size: 37; **DPF**(25): 8 15 17 18 26 30 35 37 39 51 57 64 71 72 75 91 92 119 122 135 136 137 140 145 147; **LSDPF**(3): 43 85 132; **PDF**(7): 7 42 47 62 126 131 148; **VLF**(2): 29 79;

Table 3

Members of factions

Faction title	Faction members
DPF	8 15 17 18 26 30 35 37 39 51 57 64 69 71 72 75 91 92 119 122 127 135 136 137 140 145 147
LCSF	13 21 22 60 61 81 99 123
LF	3 14 23 32 38 41 54 65 118 121 124 141
LSDPF	5 11 16 19 43 49 50 53 55 70 77 82 83 85 90 100 103 106 107 108 110 120 132 133 142 143
MG	40 112 144
NSF	44 52 74 84 105 109 113 117 128
PDF	148 7 42 47 62 67 78 97 126 131 146
TSF	2 4 10 24 25 27 28 31 34 45 46 48 58 59 63 66 88 96 98 104 115 116 130 138 139
TTLDF	1 6 33 56 68 86 129 134
VLF	9 12 29 73 76 79 89 93 95 101 111 114 125

5. Results of Multidimensional Scaling

Multidimensional scaling (MDS) is a technique for exploratory analysis of multidimensional data widely usable in different applications (Borg and Groenen, 2005; Cox and Cox, 2001). By means of this technique the set of considered objects is mapped to the low dimensionality space of images where objects are represented by points, and dissimilarities are represented by interpoint distances. Frequently a two dimensional space of images is considered aiming to visualize the results of MDS. Formally MDS does not produce partitioning of the considered set of objects. However, when objects are exposed for visual analysis they can be not only heuristically clustered but also analyzed from different angles, e.g., how compact are clusters, how they are located with respect to each other, etc. Figuratively speaking, MDS extends heuristic abilities of human experts to multidimensional spaces where they naturally do not work.

Let us give a short formulation of the problem. The dissimilarity between pairs of N objects is given by the matrix $D = (d(k, m))$, $k, m = 1, \ldots, N$, and it is supposed that $d(k, m) = d(m, k)$. The points $z_i \in R^r$, $i = 1, \ldots, N$, in an r-dimensional space of images should be found whose interpoint distances fit the given dissimilarities. Different measures of accuracy of fit can be chosen defining different images of the considered set of objects. In the case the objects are points in a high dimensional vector space such images can be interpreted as different nonlinear projections of the set of points in high dimensional space of originals to a lower dimensionality space of images. The problem of construction of images of the considered objects is reduced to minimization of an accuracy of fit criterion, e.g., of the most frequently used least squares *STRESS* function

$$STRESS(Z) = \sum_{i<j} w_{ij} \big(\delta_{ij}(Z) - d(i,j)\big)^2, \tag{7}$$

where $Z = (z_{11}, \ldots, x_{N1}, x_{12}, \ldots, x_{Nr})^T$ is vector of coordinates of images; $\delta_{ij}(X)$ denotes the distance between the ith and jth points in image space z_i and z_j; it is supposed that the weights are positive: $w_{ij} > 0$, $i, j = 1, \ldots, n$.

Since different distances $\delta_{ij}(Z)$ can be defined, the formula (7) defines a class of accuracy criteria. To define a particular criterion a norm in $\mathbb{R}^r$ should be chosen implying the particular formula for calculating distances $\delta_{ij}(Z)$. The most frequently used distances are Euclidean. However, MDS with other Minkowski distances in space of images can be even more informative than MDS with Euclidean distances. Results of MDS with different, e.g., city block and Euclidean, distances can be useful to grasp different properties of the considered objects as discussed in (Žilinskas and Žilinskas, 2007). For the algorithmic aspects of implementation of MDS algorithms with city block distances we refer to (Žilinskas and Žilinskas, 2007). Multidimensional scaling of voting data have been performed using both distances in the space of images: Euclidean distances were used in case of $d_2(\cdot, \cdot)$ dissimilarities in voting space, and city block distances were used in case of $d_1(\cdot, \cdot)$ dissimilarities. The results of MDS with two dimensional space of images are presented in Figs. 3 and 4. Let us note that adding of one or several points to the images can be done without of repeated solution of augmented optimization problem (Bernataviciene *et al.*, 2007).
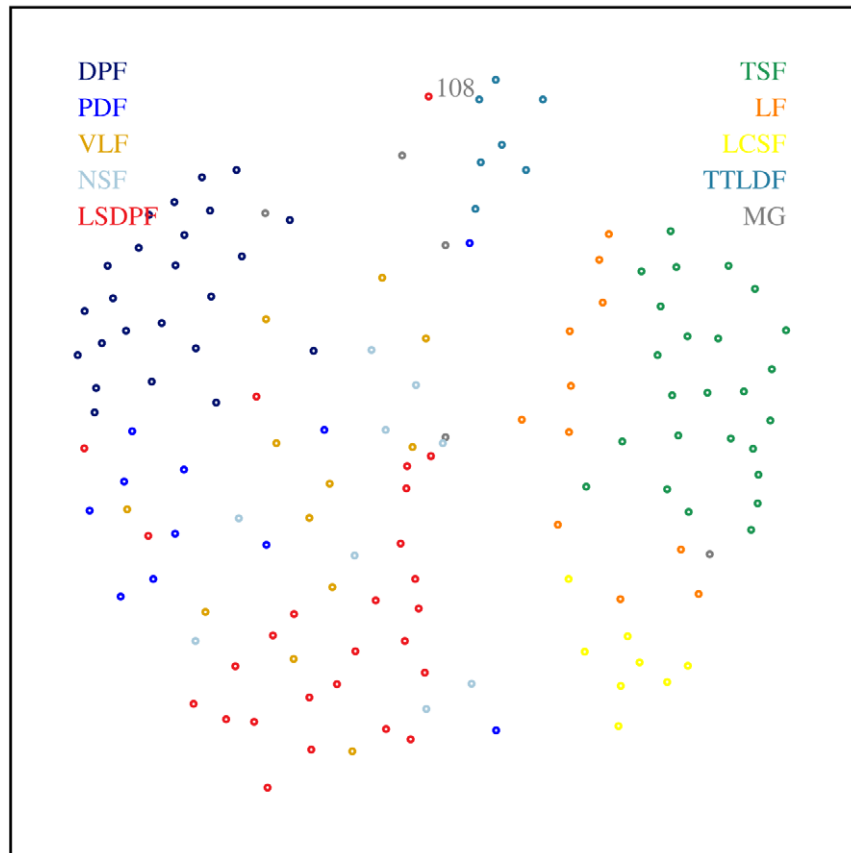
Fig. 3. Visualization of voting results using Euclidean distances.

6. Comparative Analysis

Both considered methods can be described as exploratory methods since they do not produce uniquely interpretable quantitative results. However, they produce results which can be interpreted heuristically and which induce ideas for future qualitative analysis.

The results of cluster analysis show that voting results first of all cluster according to membership of MPs to party blocks which can be specified as position and opposition. Taking into account smaller dissimilarity, more peculiar structure can be grasped, e.g., dissimilarity between groups of factions within position and within opposition is quite similar. The latter conclusion follows from closeness of second branching thresholds in dendrograms (see Fig. 1 and Fig. 2).

Discussing dissimilarities between clusters, however, it should be taken into account that furthest neighbor distance (e.g., given on vertical axis of Fig. 1 and Fig. 2) not necessary well represents heuristic concept of dissimilarity albeit it is suitable from the algorithmic point of view. The dendrogram correctly shows only pairwise distances between
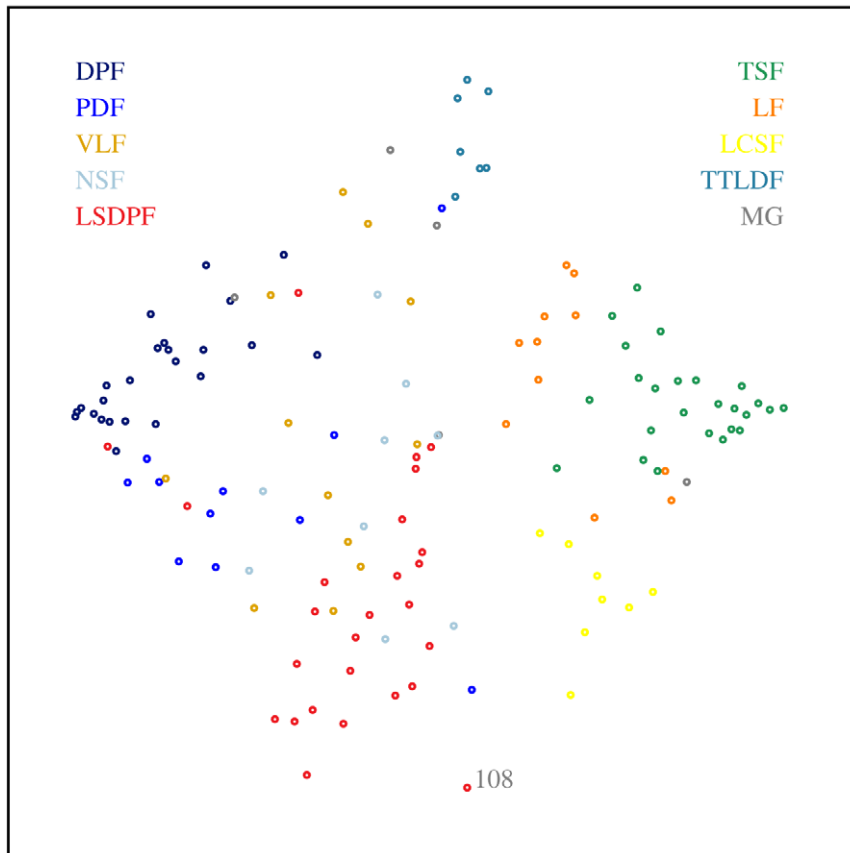
Fig. 4. Visualization of voting results using city block distances.

clusters but is not very helpful in understanding mutual disposition of larger number of clusters. However, the order of clusters merging shows the precedence of the relations among them. Summarizing, we can conclude that clustering algorithms offer a modest support in analysis of voting data; to enhance their applicability they should be augmented with visualization algorithms and algorithms of multidimensional statistical analysis.

MDS algorithms present the voting data in the form of a picture prone to heuristic analysis. The clusters indicated by means of clustering method described above are clearly visible in Fig 3 and Fig 4, e.g., position factions at the left part of the figures and opposition factions at the right part of the figures. Let us mention that the figures are invariant with respect to translation and some rotations; here they are rotated to place TSF (who declares its rightism the most strongly) at the right side of the picture. We discuss the possibility of indication of left/right in political science sense below.

These figures give rather fine grained structure of data set, e.g., the well known strongest opposition between DPF and TSF is clearly highlighted. Especially interesting is to see the leading position and opposition factions (LSDPF and TSF) with the faction

naming itself Liberal and Center Faction between them separated from others with main diagonal.

We have mentioned that the pictures of Fig. 3 and Fig. 4 are rotated rather arbitrary. It seems desirable to chose orientation of such pictures in such a way that the axes could be named sensibly. For example, in some applications choice of axes Left-Right/Libertarian-Authoritarian, North-South/West-East can be helpful in interpretation of the results. However, in our study such interpretation does not have a justification. Moreover, it is even not clear if in our case two dimensional representation of data can catch such subtle properties as Left-Right oriented politic.

Assessment of dimensionality of voting space is an active research field (Poole, 2005). A possibility to assess the actual dimensionality of data space is important advantage of MDS. This can be done by means of successive mapping of data set into image spaces of growing dimensionality and observing the decrease of mapping error. In Fig. 5 dependence of a relative visualization error (see (Žilinskas and Žilinskas, 2007)) on dimensionality of the image space is presented. Rather slow decrease of this error shows that explanation of voting results in the considered case by means of two political concepts would be superficial.

Important advantage of MDS is visual representation of voting data, including, e.g., the possibility of lightening of voting of an individual MP. This advantage is especially important for monitoring political behavior of parties and individual MPs. Wider implementation of such a method would be helpful to people for more objective assessment of candidates in current election of the parliament.

For political scientist the usefulness of MDS would be enhanced if the algorithm would be augmented with statistical analysis enabling interactive calculation of different statistical characteristics of voting of chosen subsets of MPs. Besides of such a technically simple enhancement, applicability of MDS in analysis of voting data would became more attractive if dynamics of voting results would be visualized; solution of this problem is indeed challenging.
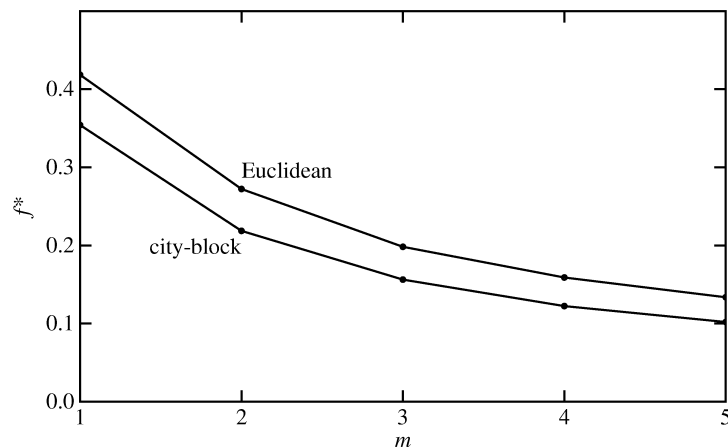

Fig. 5. Dependence of mapping error on dimensionality of the space of images.

Notice, how behavior of the MP number 108 illustrates *faction hopping*[2] and dependency on the chosen dissimilarity measure. When city block distance is chosen, MP 108 appears at the bottom of Fig. 3 and is clusterized together with the left wing parties Table 2, while in the case of Euclidean distances MP 108 is at the top of Fig. 4 and is clusterized with the right wing.

7. Conclusions

- Both, MDS and hierarchical, clusterings produce grouping of MPs well explainable in general political context.
- MDS and hierarchical clustering supplement each other, and they could be used as essential elements in development of a special toolbox for quantitative analysis of voting data.
- Voting data visualized by means of MDS well suit for monitoring of political behavior of parties and individul MPs, and such a monitoring could be recommended for unprejudiced reporting about political activities of parliaments of new democratic states.

8. Acknowledgments

References

Bernataviciene, J., G. Dzemyda and V. Marcinkevicius (2007). Conditions for optimal efficiency of relative MDS. *Informatica*, **18**(2), 187–202.

Berthold, M., and D.J. Hand (Eds.) (1999). *Intelligent Data Analysis: An Introduction*. Springer, Secaucus, NJ, USA.

Borg, I., and P. Groenen (2005). *Modern Multidimensional Scaling*, 2nd ed. Springer, New York.

Cox, T., and M. Cox (2001). *Multidimensional Scaling*. Chapman and Hall/CRC, Boca Raton.

Hix, S., and A. Noury (2007). Government-opposition or left-right? the institutional determinants of voting in eight parliaments. In *Proceedings of the Annual Conference of the American Political Sciences Association*. Chicago, Illinois, USA.

---

[2]*Faction hopping* or *parliamentary tourism* defines behavior of an MP, who switches parties or factions, usually during the parliament term.

Hix, S., A. Noury and G. Roland (2006). Dimensions of politics in the European Parliament. *American Journal of Political Science*, **50**(2), 494–511.
  `http://personal.lse.ac.uk/hix/HixNouryRolandEPdata.HTM`
Irving, F., and J. Todd (2007). The public whip. `www.publicwhip.org.uk/`. Last visited: 2007-07-08.
Keim, D., and M. Ward (1999). Visualization, chapter 11, pp. 403–427. In M. Berthold and D.J. Hand (Eds.), 1999.
Krilavičius, T., A. Žilinskas and J. Žilinskas (2007). *On Structural Analysis of Parliamentarian Voting Date: Appendices.*
  `http://surface.lt/krilaviciust/publications/data`. Last visited: 2008-01-25.
Medelis, Ž. (2007). *iDemokratija.*
  `www.idemokratija.info`. Last visited: 2007-07-08.
Poole, K.T. (2005). *Spatial Models of Parliamentary Voting*. Cambridge University Press, New York.
Seimas2007 (2007). *Seimas of the Republic of Lithuania.*
  `http://www3.lrs.lt`. Last visited: 2007-01-12.
StatToolbox (2007). *Statistics Toolbox 6: User guide, Matlab.*
  `http://www.mathworks.com/access/helpdesk/help/pdf_doc/stats/stats.pdf`
Theodoridis, S., and K. Koutroumbas (2006). *Pattern Recognition*, 3rd ed. Academic Press, Inc., Orlando, FL, USA.
Žilinskas, A., and J. Žilinskas (2006). Parallel hybrid algorithm for global optimization of problems occurring in MDS-based visualization. *International Journal Computers and Mathematics with Applications*, **52**(1–2), 211–224.
Žilinskas, A., and J. Žilinskas (2007). Two level minimization in multidimansional scaling. *Journal of Global Optimization*, **38**, 581–596.

T. Krilavičius is a post doctoral student at the Informatics faculty of Vytautas Magnus University. His research interests are rigorous analysis and visualization of political processes, formal specification, analysis and simulation of hybrid systems (behavioural hybrid process calculus, hybrid automaton).

A. Žilinskas is the chairman of Optimization Sector and a principal researcher at the Institute of Mathematics and Informatics, Lithuania. His scientific interests are statistical theory of global optimization, optimization based modelling and design, and analysis of multidimensional data by means of visualization. His research is oriented to develop statistical models of global optimization, implement and investigate the corresponding algorithms, and apply them to practical problems. He is a member of IEEE including Computer Society and Computational Intelligence Society, of American Mathematical Society, and of IFIP working group Optimization Based Computer Aided Modelling and Design as well as member of editorial boards of Journal of Global Optimization, Control and Cybernetics, Informatica, The Open Cybernetics and Systemics Journal, and a new journal in preparation International Journal of Grid and High Performance Computing.

Struktūrinė seimo narių balsavimo analizė

Tomas KRILAVIČIUS, Antanas ŽILINSKAS

Tiriamas populiarių kiekybinių metodų tinkamumas analizuoti struktūrines savybes duomenų, atspindinčių Lietuvoje vykstančius fundamentalius politinius procesus. Analizuojami Lietuvos Seimo balsavimo duomenys. Lyginami du dažniausiai taikomi struktūrinės analizės metodai: klasterizavimas ir daugiamatės skalės. Pateikiamos techninės išvados, kurios gali būti naudingos tolimesniam šių metodų adaptavimui ir taikymui politiką atspindinčių duomenų analizėje.