

CATEGORY SELECTION AND CLASSIFICATION BASED ON CORRESPONDENCE COORDINATES

OTTÓ HAJDU¹

The paper presents the description and an application of the explorative multivariate technique known as multiple correspondence analysis of an indicator matrix. Correspondence coordinates have been used to reveal relevant categories of economic organizations in connection with their financial bankruptcy. Illustrative calculations are based on data from balance sheets of Hungarian enterprises. The aim of the paper is twofold. On the one hand, it seeks correspondences among categories of the variables investigated. On the other hand, based on the relevant categories it discriminates the two groups of active and bankrupt firms and classifies an additional supplementary category of firms to one of them. This third category is the group of those who are still currently active but already affected by bankruptcy proceedings. Finally, an individual firm is also predicted. To clarify the meaning of the correspondence coordinates a detailed explanation of their theory is provided.

KEYWORDS: Explorative multivariate techniques; Categorical data analysis; Multiway contingency tables.

Correspondence analysis is an exploratory multivariate technique that converts a data matrix of non-negative numbers (usually frequency table) into a graphical display in which rows and columns are depicted as points. By comparing row and column proportions in a two- or multiway table it provides a method for visually interpreting multivariate categorical data. Especially, displaying row and column profiles as points in a two dimensional subspace we can discuss the structure of association between the row and column categories.

Simple correspondence analysis (CA) involves two categorical variables and the graphical display of the corresponding two-way contingency table. Mathematically, CA decomposes the Pearson- χ^2 measure of association for the table into components to underlie dimensions of heterogeneity between rows or columns. This is done in a manner similar to that of variance decomposition by principal component analysis for continuous data. On the other hand, CA simultaneously assigns a scale to rows and a separate scale to columns so as to maximize the correlation between the resulting pairs of variables.

For multiple correspondence analysis (MCA) the latter concept is the more appropriate. MCA is an extension of CA to the case of three or more categorical variables. It is

¹ Associate professor of the Budapest University of Technology and Economic Sciences.

characterized by similar graphical displays in which either the categories of the variables or the individual cases themselves can be represented as points. MCA resembles a principal component analysis for categorical variables.

Using MCA the main purpose of this paper is to select appropriate predictor categories associated with a financial bankruptcy in an average sense based on some economic data of Hungarian enterprises. Besides, taking outcomes of the relevant predictor variables (such as type of activity, legal form, the level of profitability etc.) into account, we illustrate how to classify an additional, currently active organization whether it seems similar (or not) to those who had finished their activities due to a bankruptcy. Although, category selection from more than two scales prefers using MCA, a brief but detailed overview of the general theory of correspondence axes is necessary because MCA is merely an application of CA carried out on a special contingency table with individual cases as rows and categories (expressed by dummy variables) as columns. This kind of the data set is called indicator matrix. In order to serve a guide to the correct interpretation of the MCA results we also focus on the specific considerations that must be taken into account due to the special form of the indicator matrix.

PROPERTIES OF THE CORRESPONDENCE COORDINATES

The measure of association for a two-way contingency table of n observations is completely determined by the pattern of the relative frequencies in the table proportional to the grand total n . Considering the $p_{ij} = f_{ij}/n$ joint relative frequency of row i and column j CA basically analyses the correspondence matrix with elements p_{ij} (see Table 1) where f_{ij} is the observed frequency in the j^{th} column of row i .

Table 1

<i>Correspondence table</i>						
Category	Column				Total: mass of the row	
	1	...	j	...		J
Row 1	p_{11}		p_{1j}		p_{1J}	s_1
⋮						⋮
Row i	p_{i1}		$p_{ij} = f_{ij}/n$		p_{iJ}	s_i
⋮						⋮
Row I	p_{I1}		p_{Ij}		p_{IJ}	s_I
<i>Total: mass of the column</i>	o_1		o_j		o_J	1

The row total s_i and the column total o_j are the relative marginal (unconditional) frequencies termed masses expressed also as percentages of the grand total n . Considering the conditional set of relative frequencies within a row category we use the term *row profile* and respectively within a column the term *column profile* (see Table 2 and Table 3). The row profiles are treated as points in the J -dimensional space spanned by the columns while column profiles are points in the I -dimensional space spanned by the rows. Hence, the row and column profiles constitute two clouds of points in respective J - and I -

dimensional spaces. The associated masses of the axes are included in and denoted by the vectors \mathbf{s} and \mathbf{o} and assigned to these axes as weights.

Table 2

Row profiles: $[\mathbf{R}]_{ij}=s_{ij}$

Row profile	Axis					Total
	1	...	j	...	J	
1	s_{11}		s_{1j}		s_{1J}	1
\vdots						\vdots
i	s_{i1}		$s_{ij}=p_{ij}/s_i$		s_{iJ}	1
\vdots						\vdots
I	s_{I1}		s_{IJ}		s_{IJ}	1
<i>Centroid (mass)</i>	o_1		o_j		o_J	1

Table 3

Column profiles: $[\mathbf{C}]_{ij}=o_{ij}$

Axis	Column profile					Centroid (mass)
	1	...	j	...	J	
1	o_{11}		o_{1j}		o_{1J}	s_1
\vdots						\vdots
i	o_{i1}		$o_{ij}=p_{ij}/o_j$		o_{iJ}	s_i
\vdots						\vdots
I	o_{I1}		o_{IJ}		o_{IJ}	s_I
<i>Total</i>	1		1		1	1

It is obvious, that the row and column profiles are closely related to each other and to the elements of the correspondence table as follows:

$$p_{ij} = s_i \cdot s_{ij} = o_j \cdot o_{ij} \quad /1/$$

Taking the summations of both sides in /1/ it is apparent based on Table 1 that masses s_i and o_j are the weighted averages of the column and row profiles respectively, using the other set of masses as weights:

$$s_i = \sum_{j=1}^J o_j o_{ij}, \quad o_j = \sum_{i=1}^I s_i s_{ij}.$$

Hence, the o_j masses of the columns constitute the centroid of the row profiles and the s_i masses of the rows the centroid of the column profiles. Then, the independence (lack of association) between the row and column clouds is defined as row profiles identical to each other and hence to their centroid too. Necessarily, when the lack of association occurs in the contingency table the column profiles are also identical to each other. In other words, a non-zero variation of the points in a cloud indicates a lack of independence.

Comparing then simply the row profiles with respect to the columns as axes or comparing the column profiles with respect to the rows as axes reveals the nature of association between rows and columns.

However, when the number of the rows or the columns or both is too large it is difficult to identify similarities and dissimilarities by simply scanning the row and column percentages. Then, information on association involved in the contingency table can be summarized briefly by the well-known Pearson- χ^2 measure which hereafter will be termed the *total inertia*:

$$INR = \sum_{i=1}^I \sum_{j=1}^J \frac{(p_{ij} - s_i o_j)^2}{s_i o_j} = \sum_{i=1}^I \sum_{j=1}^J g_{ij}^2, \quad /2/$$

where $s_i o_j$ is the expected relative frequency of cell (i, j) for the case of independence and

$$g_{ij} = \frac{p_{ij} - s_i o_j}{\sqrt{s_i o_j}} \quad /3/$$

is the standardized correspondence frequency. Value of g_{ij} that deviates markedly from zero indicates a positive or negative association between row i and column j . Based on equation /1/ the total inertia can be expressed as a weighted multidimensional dispersion measure considering either the row or the column profiles as points:

$$INR = \sum_{i=1}^I s_i \sum_{j=1}^J \frac{1}{o_j} (s_{ij} - o_j)^2 = \sum_{i=1}^I s_i \sum_{j=1}^J \frac{s_{ijc}^2}{o_j}, \quad /4/$$

$$INR = \sum_{j=1}^J o_j \sum_{i=1}^I \frac{1}{s_i} (o_{ij} - s_i)^2 = \sum_{j=1}^J o_j \sum_{i=1}^I \frac{o_{ijc}^2}{s_i} \quad /5/$$

where $s_{ijc} = s_{ij} - o_j$ and $o_{ijc} = o_{ij} - s_i$ are the centered row and column profiles respectively. In this context *INR* is a multivariate extension of variance defined as the weighted average of the squared deviates from the respective centroid. It is to be noted, that the centroid of a centered profile is always the origin.

In CA, instead of comparing the rows using directly the centered profiles we create a smaller number of coordinates. These coordinates are computed so that each successive ($k=1, 2, \dots, K$) coordinate axis accounts for a decreasing portion of the total inertia. The first coordinate accounts for the largest part, the second for the next largest part, and so on. The first coordinate or the first two coordinates often account for the major part about 80-90 percent or more. When these first two coordinates explain most of the inertia we can summarize each row of them instead of the original row percentages. This permits almost all of the information to be presented in a one- or two-dimensional plot. The same argument holds for analyzing the column pattern. The centered row profiles are replaced by CA coordinates x presented in the matrix $\mathbf{X}_{(I,K)}$ and the centered transposed column profiles are replaced by CA coordinates y presented in the matrix $\mathbf{Y}_{(J,K)}$. The maximum number of the new axes is $K = \min\{I-1, J-1\}$ because the relative frequencies within a profile always sum up to 1. This is shown in Table 4 and Table 5.

Table 4

Centered row profiles and their correspondence coordinates

Row profile	Centered profile: $[S]_{j\bar{c}}=s_{ijc}$					Row CA coordinate: X				
1	s_{11c}	...	s_{1jc}	...	s_{1Jc}	x_{11}	...	x_{1k}	...	x_{1K}
\vdots						\vdots				\vdots
i	s_{i1c}		s_{ijc}		s_{iJc}	x_{i1}		x_{ik}		x_{iK}
\vdots						\vdots				\vdots
I	s_{I1c}		s_{IJc}		s_{IJc}	x_{I1}		x_{Ik}		x_{IK}
<i>Centroid</i>	0		0		0	0		0		0

Table 5

Transposed centered column profiles and their correspondence coordinates

Column profile	Centered profile: $[O^T]_{j\bar{c}}=o_{jic}$					Column CA coordinate: Y				
1	o_{11c}	...	o_{1ic}	...	o_{1Jc}	y_{11}	...	y_{1k}	...	y_{1K}
\vdots						\vdots				\vdots
j	o_{j1c}		o_{jic}		o_{jJc}	y_{j1}		y_{jk}		y_{jK}
\vdots						\vdots				\vdots
J	o_{J1c}		o_{Jic}		o_{JJc}	y_{J1}		y_{Jk}		y_{JK}
<i>Centroid</i>	0		0		0	0		0		0

The computed CA coordinates (over all extractable dimensions) are required to leave the inertia of a point unchanged:

$$INR(s_i) = s_i \sum_{k=1}^K x_{ik}^2 = \sum_{j=1}^J s_i \frac{s_{ijc}^2}{o_j}, \quad INR(o_j) = o_j \sum_{k=1}^K y_{jk}^2 = \sum_{i=1}^I o_j \frac{o_{jic}^2}{s_i}, \quad /6/$$

where

$$\sum_{k=1}^K s_i x_{ik} = 0, \quad \sum_{k=1}^K o_j y_{jk} = 0. \quad /7/$$

Equation /6/ says that the sum of the squared CA coordinates preserves the information entirely. Consequently, using equations /4/ and /6/ the total inertia also remains unchanged:

$$INR = \sum_{i=1}^I INR(s_i) = \sum_{j=1}^J INR(o_j).$$

Along with a separated CA axis k the measure of inertia reduces to variance so that the variance of the rows and the variance of the columns are equal:

$$Var(x | k) = \sum_{i=1}^I s_i x_{ik}^2 = Var(y | k) = \sum_{j=1}^J o_j y_{jk}^2 = Var(k),$$

where $Var(k)$ is the inertia of the CA axis k . This property will be clarified by equation /16/. The spread of the inertia is illustrated in Table 6. It is obvious that the total inertia is partitioned by the CA axes as follows: $INR = \sum_{k=1}^K Var(k)$.

Table 6

The structure of inertia

Point	Correspondence axis					Total
	1.	...	k	K .	
Row 1	$s_1x_{11}^2$		$s_1x_{1k}^2$		$s_1x_{1K}^2$	$INR(s_1)$
⋮						⋮
Row i	$s_ix_{i1}^2$		$s_ix_{ik}^2$		$s_ix_{iK}^2$	$INR(s_i)$
⋮						⋮
Row I	$s_Ix_{I1}^2$		$s_Ix_{Ik}^2$		$s_Ix_{IK}^2$	$INR(s_I)$
Total	$Var(1)$		$Var(k)$		$Var(K)$	INR
Column 1	$o_1y_{11}^2$		$o_1y_{1k}^2$		$o_1y_{1K}^2$	$INR(o_1)$
⋮						⋮
Column j	$o_jy_{j1}^2$		$o_jy_{jk}^2$		$o_jy_{jK}^2$	$INR(o_j)$
⋮						⋮
Column J	$o_Jy_{J1}^2$		$o_Jy_{Jk}^2$		$o_Jy_{JK}^2$	$INR(o_J)$

In order to calculate CA coordinates let us define the diagonal matrices $\mathbf{D}_s = \langle s_1, \dots, s_I \rangle$, $\mathbf{D}_o = \langle o_1, \dots, o_J \rangle$, $\mathbf{D}_\mu = \langle \mu_1, \dots, \mu_K \rangle$ and the matrix $\mathbf{G}_{(I,J)}$ with the g_{ij} standardized correspondence frequencies as its elements. At this stage based on equations /1/ and /3/ we rewrite matrix \mathbf{G} as $\mathbf{G} = \mathbf{D}_s^{1/2} \mathbf{S} \mathbf{D}_o^{-1/2} = \mathbf{D}_s^{-1/2} \mathbf{O} \mathbf{D}_o^{1/2}$ and then take its $\mathbf{G} = \mathbf{U} \mathbf{D}_\mu \mathbf{V}^T$ so-called 'Singular Value Decomposition' (SVD)². By definition of SVD this yields the following equation:

$$\mathbf{G} = \mathbf{D}_s^{1/2} \mathbf{S} \mathbf{D}_o^{-1/2} = \mathbf{D}_s^{-1/2} \mathbf{O} \mathbf{D}_o^{1/2} = \mathbf{U} \mathbf{D}_\mu \mathbf{V}^T, \quad /8/$$

where $\mu_1, \mu_2, \dots, \mu_K$ are the singular values, the columns of the matrix $\mathbf{U}_{(I,K)}$ are the left singular vectors and the columns of the matrix $\mathbf{V}_{(K,K)}$ are the right singular vectors of \mathbf{G} satisfying the orthonormality requirement of $\mathbf{U}^T \mathbf{U} = \mathbf{V}^T \mathbf{V} = \mathbf{I}$. (\mathbf{I} stands for the \mathbf{I}_K identity matrix.) The columns of \mathbf{U} define the principal axes of the column cloud and the columns of \mathbf{V} define the principal axes of the row cloud of \mathbf{G} . Now, the \mathbf{X} and \mathbf{Y} CA coordinates of our interest are defined as the principal coordinates with respect to the principal axes of \mathbf{S} and \mathbf{O} respectively. From /8/ the weighted SVD of \mathbf{S} and \mathbf{O}^T yields:

$$\mathbf{S} = \left(\mathbf{D}_s^{-1/2} \mathbf{U} \mathbf{D}_\mu \right) \left(\mathbf{D}_o^{1/2} \mathbf{V} \right)^T = \mathbf{X} \left(\mathbf{D}_o^{1/2} \mathbf{V} \right)^T \quad /9/$$

² For more details see e.g. Greenacre (1984).

$$\mathbf{O}^T = \left(\mathbf{D}_o^{-1/2} \mathbf{V} \mathbf{D}_\mu \right) \left(\mathbf{D}_s^{1/2} \mathbf{U} \right)^T = \mathbf{Y} \left(\mathbf{D}_s^{1/2} \mathbf{U} \right)^T, \quad /10/$$

where

$$\mathbf{X} = \mathbf{D}_s^{-1/2} \mathbf{U} \mathbf{D}_\mu = \mathbf{S} \mathbf{D}_o^{-1/2} \mathbf{V} \quad /11/$$

$$\mathbf{Y} = \mathbf{D}_o^{-1/2} \mathbf{V} \mathbf{D}_\mu = \mathbf{O}^T \mathbf{D}_s^{-1/2} \mathbf{U}. \quad /12/$$

Alternatively, the transition of the column coordinates into row coordinates is also possible:

$$\mathbf{X} = \mathbf{S} \mathbf{Y} \mathbf{D}_\mu^{-1} = \left(\mathbf{R} - \mathbf{1} \mathbf{o}^T \right) \mathbf{Y} \mathbf{D}_\mu^{-1} = \mathbf{R} \mathbf{Y} \mathbf{D}_\mu^{-1}, \quad /13/$$

where, conversely, the transition of the row coordinates into column coordinates in a similar manner is given by:

$$\mathbf{Y} = \mathbf{O}^T \mathbf{X} \mathbf{D}_\mu^{-1} = \left(\mathbf{C}^T - \mathbf{1} \mathbf{s}^T \right) \mathbf{X} \mathbf{D}_\mu^{-1} = \mathbf{C}^T \mathbf{X} \mathbf{D}_\mu^{-1}, \quad /14/$$

where $\mathbf{s} = \text{diag} \mathbf{D}_s$, $\mathbf{o} = \text{diag} \mathbf{D}_o$ and recall that based on equation (7) the centroid of the CA coordinates is the origin that is $\mathbf{o}^T \mathbf{Y} = \mathbf{0}^T$ and $\mathbf{s}^T \mathbf{X} = \mathbf{0}^T$. Writing in more details:

$$x_{ik} = \sum_{j=1}^J \frac{s_{ij} y_{jk}}{\mu_k}, \quad y_{jk} = \sum_{i=1}^I \frac{o_{ij} x_{ik}}{\mu_k}. \quad /15/$$

It is to be noted, that each row coordinate is a weighted average of the standardized column coordinates with row profile elements as weights and conversely. It is obvious that x_{ik} and y_{jk} tend to be close to each other when column j has a large s_{ij} proportion in the row profile i or row i has a large o_{ij} proportion in the column profile j . In this case a large row coordinate on the CA axis k necessarily yields also a large column coordinate on the same axis. Thus, the CA row and column coordinates can be thought of as the result of a dual scaling of row and column scales. The pair of sets of coordinates x_{i1} and y_{j1} provide one dual scaling with standard deviation μ_1 , along one dimension while x_{i2} and y_{j2} provide another dual scaling in an orthogonal dimension with standard deviation μ_2 etc. A very important role of the transition formulas /13/ and /14/ in CA to add supplementary points (either rows or columns) to the CA plots. In other words, transition formulas enable us to predict row or column profiles that are omitted from the actual computation of the CA coordinates.

From equations /11/ and /12/ follows that the covariance matrix \mathbf{Cov}_{xx} of coordinates \mathbf{X} and the covariance matrix \mathbf{Cov}_{yy} of the coordinates \mathbf{Y} are identical and diagonal with the squared singular values in the main diagonal:

$$\mathbf{Cov}_{xx} = \mathbf{X}^T \mathbf{D}_s \mathbf{X} = \mathbf{Cov}_{yy} = \mathbf{Y}^T \mathbf{D}_o \mathbf{Y} = \mathbf{D}_\mu^2 = \left\langle \mu_1^2, \mu_2^2, \dots, \mu_k^2 \right\rangle, \quad /16/$$

where $\mu_k^2 = \text{Var}(x | k) = \text{Var}(y | k)$ termed the 'principal inertia'.

It must be emphasized at this stage that it is always possible to analyze a higher-dimensional table in a two-way form. Then, a row refers to a combination of the levels of two or more variables and a column refers to a combined category of another set of variables. This method of forming two new combined variables is called ‘stacking’.

MEASURING GOODNESS OF FIT

Extracting exactly the first $m < K$ leading CA axes (typically one or two) the question arises that how well the points are represented in the reduced lower dimensional subspace. Information involved in the CA coordinates is summarized in the following goodness of fit measures.

a) Inertia explained:

$$IE(m) = \frac{\sum_{k=1}^m \mu_k^2}{\sum_{k=1}^K \mu_k^2}.$$

This measure tells us that the first m axes account for the $IE(m)$ percentage of the total inertia. Supposed that almost all of the inertia is accounted for by the first two axes, it indicates that a two-dimensional representation of the rows and columns is very accurate.

b) The quality of a point:

$$QLT_i(m) = \frac{\sum_{k=1}^m x_{ik}^2}{\sum_{k=1}^K x_{ik}^2}, \quad QLT_j(m) = \frac{\sum_{k=1}^m y_{jk}^2}{\sum_{k=1}^K y_{jk}^2}.$$

This measure indicates the contribution of the first m principal axes to the inertia of row i or column j respectively. A low quality implies that the point considered (row or column) lies outside of the m dimensional plane.

c) Contribution to the inertia of an axis:

$$CTR_{ik} = \frac{s_i x_{ik}^2}{\mu_k^2}, \quad CTR_{jk} = \frac{o_j y_{jk}^2}{\mu_k^2},$$

where $\sum_i CTR_{ik} = \sum_j CTR_{jk} = 1$. This measure reports the relative contribution of row i or column j to the inertia (variance in this case) of the k^{th} CA axis.

d) Squared correlation:

$$COR2_{ik} = \frac{s_i x_{ik}^2}{INR(s_i)}, \quad COR2_{jk} = \frac{o_j y_{jk}^2}{INR(o_j)}.$$

It reports the contribution of axis k to the inertia of row i and column j respectively. A low $COR2$ implies that the point is not well represented in that dimension.

Apparently, QLT and $COR2$ are independent on the marginal proportions (masses) whilst CTR depends on the masses.

When results are poor regarding the goodness of fit measures it is suggested to confirm the results by collapsing or deleting certain table categories. Checking results within a stratum (i.e. a selected level) of an additional variable could also be meaningful.

As an immediate consequence of /11/ and /12/ we have the following formula for reconstituting the elements p_{ij} of the correspondence matrix \mathbf{P} :

$$\mathbf{G} = (\mathbf{D}_s^{1/2} \mathbf{X}) \mathbf{D}_\mu^{-1} (\mathbf{D}_o^{1/2} \mathbf{Y})^T,$$

that is

$$g_{ij} = \frac{p_{ij} - s_i o_j}{\sqrt{s_i o_j}} = \sqrt{s_i o_j} \sum_{k=1}^K \frac{x_{ik} y_{jk}}{\mu_k}$$

and finally the exact and then the approximate reconstitution p_{ij} value:

$$p_{ij} = s_i o_j \left(1 + \sum_{k=1}^K \frac{x_{ik} y_{jk}}{\mu_k} \right) \approx \quad /17/$$

$$\approx s_i o_j \left(1 + \sum_{k=1}^m \frac{x_{ik} y_{jk}}{\mu_k} \right). \quad /18/$$

The approximate reconstitution of the p_{ij} from the CA axes display can be used on the one hand to impute missing values in the data matrix. On the other hand, underlining the situation when $m=1$ we stress that it is not only the closeness of a row point to a column point that determines their degree of association, but also the comparison of their distances from the origin. Therefore when the product $x_{i1} y_{j1}$ is near zero, the standardized deviate g_{ij} is also near zero and consequently, the association between the i^{th} row and the j^{th} column is low.

MULTIPLE CORRESPONDENCE ANALYSIS (MCA)

When more than two discrete variables have been observed on each of the n individuals, instead of stacking variables MCA is a more appropriate tool for multiple analysis. The multiple analysis is equivalent to a simple CA carried out on the so-called indicator matrix. The rows of the indicator matrix $\mathbf{Z}_{(n,Q)}$ correspond to the $i=1,2,\dots,n$ observational units (individuals, cases) of the study, while the columns correspond to the categories of the discrete variables Z_q ($q=1,2,\dots,Q$) where Z_q has J_q categories. Thus, the matrix consists of Q sets of J columns ($J = J_1 + J_2 + \dots + J_Q$) and each row has Q ones indicating the categories into which the observational units fall. This is illustrated in Table 7.

There are nQ ones scattered throughout \mathbf{Z} , n in each submatrix \mathbf{Z}_q , otherwise the elements of \mathbf{Z} are zeros. Each row of \mathbf{Z}_q adds up to 1, and each row of \mathbf{Z} adds up to Q .

Table 7

Indicator matrix

Individual (case)	Columns of the indicator matrix \mathbf{Z} ($j=1,2,\dots,J$)												Total		
	Categories of variable: Z_1				...	Categories of variable: Z_q				...	Categories of variable: Z_Q				
	1	2	...	J_1	...	1	2	...	J_q	...	1	2		...	J_Q
1	1						1							1	$\frac{Q}{Q}$
2		1					1					1			$\frac{Q}{Q}$
\vdots															\vdots
i				1		1						1			$\frac{Q}{Q}$
\vdots															\vdots
n		1							1		1				$\frac{Q}{Q}$
Total (f_j)	f_1^1	f_2^1	...	$f_{J_1}^1$...	f_1^q	f_2^q	...	$f_{J_q}^q$...	f_1^Q	f_2^Q	...	$f_{J_Q}^Q$	$n \cdot \frac{Q}{Q}$

Interpretation of the MCA results based on the following properties of \mathbf{Z} .

1. The sum of the masses $o_j=f_j/(nQ)$ of the columns of \mathbf{Z}_q is $1/Q$ for all $q=1,2,\dots,Q$. Thus each discrete variable q receives the same mass, which is distributed over the $1,2,\dots,J_q$ categories according to the frequencies f^q of responses.

2. The centroid of the $o_{ij}=(1/f_j)=1/(n \cdot Q \cdot o_j)$ column profiles of \mathbf{Z}_q is at the centroid of all the column profiles. Thus each sub cloud of categories is balanced at the origin of the display. Further, each row mass is $s_i=Q/(n \cdot Q)=1/n$ and each row profile element is $s_{ij}=1/Q$.

3. The inertia shared by a single cell of row i and column j (from equations /1/ and /2/ is

$$INR(i,j) = s_{ij}o_{ij} - 2p_{ij} + s_i o_j$$

hence, the inertia shared by a single column j is

$$\sum_i INR(i,j) = INR(j) = f_j s_{ij} o_{ij} - 2f_j p_{ij} + n s_i o_j = 1/Q \cdot o_j.$$

The inertia contributed by a category increases as the response to this category decreases, with an upper bound of $1/Q$.

4. The inertia of the column profiles of \mathbf{Z}_q is:

$$INR(q) = \sum_{j_q=1}^{J_q} INR(j_q) = \frac{J_q}{Q} - \frac{1}{Q}.$$

The inertia contributed by a discrete variable increases linearly with the number of the response categories.

5. The total inertia of the column profiles (and of the row profiles) is:

$$INR = \sum_{q=1}^Q INR(q) = \frac{J}{Q} - 1.$$

6. The number of non-trivial dimensions with positive inertia is at most $J-Q$.

7. The row profiles lie at the equal-weighted barycentre of the column profiles representing their responses, up to a re-scaling by the inverse square root of the principal inertias along the respective principal axis.

8. The n row profiles are vectors originally in the J -dimensional space, but they occur at only $J_1 \times J_2 \times \dots \times J_Q$ distinct positions.

9. In general, only principal inertias above the value $1/Q$ are ‘interesting’ and it is clear that a rather pessimistic impression of the quality of a display is obtained by the usual percentages of inertia. Especially, when the J_q categories are derived from segmenting the range of a continuous variable then we have the undesirable result that the usual percentages tend to zero, even on the major dimensions, as the subdivisions are made finer and finer. Considering an indicator matrix with $J_1 \times J_2 \times \dots \times J_Q$ rows, one row for each of the possible responses to the Q variables, then the J - Q principal inertias are all $1/Q$. This is the justification for taking $1/Q$ as a ‘baseline’ value for the principal inertias. Considering the goodness of fit measures of an MCA application, it is apparent that it is not the magnitude of their values rather the rank positions that are informative to make a selection of influential categories.

10. The standard coordinates (of the rows or the columns) in the CA analysis of $\mathbf{Z}^T \mathbf{Z}$ are identical to the standard coordinates of the columns in the CA analysis of \mathbf{Z} . The positive semidefinite (J, J) order symmetric matrix $\mathbf{Z}^T \mathbf{Z}$ is called the Burt matrix. This property follows directly from the transition formula /13/ which can be written as:

$$\mathbf{C}^T \mathbf{X} \mathbf{D}_\mu^{-1} = \mathbf{C}^T \mathbf{R} \mathbf{Y} \mathbf{D}_\mu^{-2} \quad /19/$$

and using /14/ the column coordinates \mathbf{Y} of \mathbf{Z} satisfy the following eigen-equation:

$$\mathbf{Y} = \mathbf{C}^T \mathbf{R} \mathbf{Y} \mathbf{D}_\mu^{-2} \quad /20/$$

Now, the row profile matrix \mathbf{R} is simply $(1/Q)\mathbf{Z}$, while the column profile matrix is $\mathbf{C}^T = (nQ\mathbf{D}_o)^{-1} \mathbf{Z}^T$. Since the column masses \mathbf{D}_o of \mathbf{B} are identical to those of \mathbf{Z} , equation /20/ can be written as

$$\mathbf{Y} = \left(\frac{1}{nQ^2} (\mathbf{D}_o^{Burt})^{-1} \mathbf{Z}^T \mathbf{Z} \right) \mathbf{Y} \mathbf{D}_\mu^{-2} = \mathbf{R}^{Burt} \mathbf{Y} \mathbf{D}_\mu^{-2} \quad /21/$$

Because $[nQ^2 \mathbf{D}_o^{(Burt)}]^{-1} \mathbf{Z}^T \mathbf{Z}$ is the row profile matrix of the Burt matrix and the row and column coordinates of \mathbf{B} are identical, /21/ is precisely the transition formula in the analysis of the Burt matrix. Hence, $\mathbf{Y}^Z = \mathbf{Y}^B$. The principal inertias $\mu^{2(B)}$ in the analysis of the Burt matrix are the squares of those of the indicator matrix: $\mu^{2(B)} = (\mu^{2(Z)})^2$.

It is instructive at this point to compare the analysis of \mathbf{Z} with that of the Burt matrix. Using the \mathbf{Z}_q submatrices from Table 7, the Burt matrix has the following block structure:

$$\mathbf{Z}^T \mathbf{Z}_{(J,J)} = \mathbf{B} = \begin{bmatrix} \mathbf{Z}_1^T \mathbf{Z}_1 & \mathbf{Z}_1^T \mathbf{Z}_2 & \cdots & \mathbf{Z}_1^T \mathbf{Z}_Q \\ \mathbf{Z}_2^T \mathbf{Z}_1 & \mathbf{Z}_2^T \mathbf{Z}_2 & & \mathbf{Z}_2^T \mathbf{Z}_Q \\ \vdots & & \ddots & \\ \mathbf{Z}_Q^T \mathbf{Z}_1 & \mathbf{Z}_Q^T \mathbf{Z}_2 & & \mathbf{Z}_Q^T \mathbf{Z}_Q \end{bmatrix}.$$

Each ‘off-diagonal’ submatrix $\mathbf{Z}_q^T \mathbf{Z}_{q^*}$ ($q \neq q^*$) is a two-way contingency table which condenses the association between variables q and q^* across the n individuals (cases). Each ‘diagonal’ submatrix $\mathbf{Z}_q^T \mathbf{Z}_q$ is the diagonal matrix of the column sums of \mathbf{Z}_q . Because the Burt matrix is positive, semidefinite and symmetric, it is clear that its CA produces two identical sets of coordinates for the rows and columns. The only difference between the analysis of \mathbf{B} and \mathbf{Z} lies in the values of the principal inertias, which will affect the scales of the principal coordinates.

The fact that the analysis of the multivariate indicator matrix \mathbf{Z} is equivalent to that of the Burt matrix illustrates that these analyses should be regarded as joint bivariate rather than multivariate ones. The Burt matrix is the analogue of the covariance matrix of Q continuous variables, where each (J_q, J_{q^*}) submatrix is analogous to a covariance. The CA of \mathbf{Z} (or equivalently, of \mathbf{B}) does not take into account associations among more than two discrete variables but rather looks at all the two-way associations jointly. In the context of multiway contingency table analysis we consider only the second-order interactions. Thus the CA treatment of a multivariate indicator matrix \mathbf{Z} seems to be at an interface between the classical joint bivariate treatment of continuous multivariate data and the complex interaction modelling of multiway contingency tables.

PREDICTION OF FINANCIAL BANKRUPTCY

In the following analysis the prediction of financial bankruptcy is illustrated based on the data set of Hungarian corporations and unincorporated enterprises (firms hereafter).

The data set

The data characterize the years of 1998 and 1999. The variables investigated are partly categorical and partly continuous but measured also on a scale of categories. The categorical variables of interest are the status (Status) and the legal form of the firm (F) and the type of industry (I) and the region that the firm belongs to (R). The corresponding categories are as follows.

– The Status=OK, if the firm is active, Status=BRUPT if it has finished its activity due to bankruptcy proceedings and Status=PROC, if the unit is actually under bankruptcy proceedings.

– The Legal Form takes the values of $F=\{\text{COP,GP,LP,LLC,JSC}\}$ if the firm is a Co-operative (COP), a General Partnership (GP), a Limited Partnership (LP), a Limited Liability Company (LLC), a Joint Stock Company (JSC) respectively or $F=\text{Other}$ otherwise.

– The type of industry (according to the SNA classification) is indicated simply by its order number $I=\{i1, i2, \dots, i15\}$.

– The categories of Region are $R=\{\text{CHU,CTD,WTD,STD,NHU,NGP,SGP}\}$, corresponding to regions of Central Hungary (CHU), Central Transdanubia (CTD), Western Transdanubia (WTD), Southern Transdanubia (STD), Northern Hungary (NHU), Northern Great Plain (NGP) and Southern Great Plain (SGP).

Further, four continuous financial indicators are also of our interest as potential indicators of the activity moving toward bankruptcy.

The number of 'OK and BRUPT' firms together is 169 610 from which 321 is bankrupt whilst the number of PROC firms is 3682. Because PROC firms are to be classified they are omitted from the computations. Hence, the number of cases contributing to our analysis is 169 610 constituting the 169 610 rows of the indicator matrix with 45 or 43 columns depending on the current analysis. Computations were made by using the BMDP Program Package.

Plotting associations

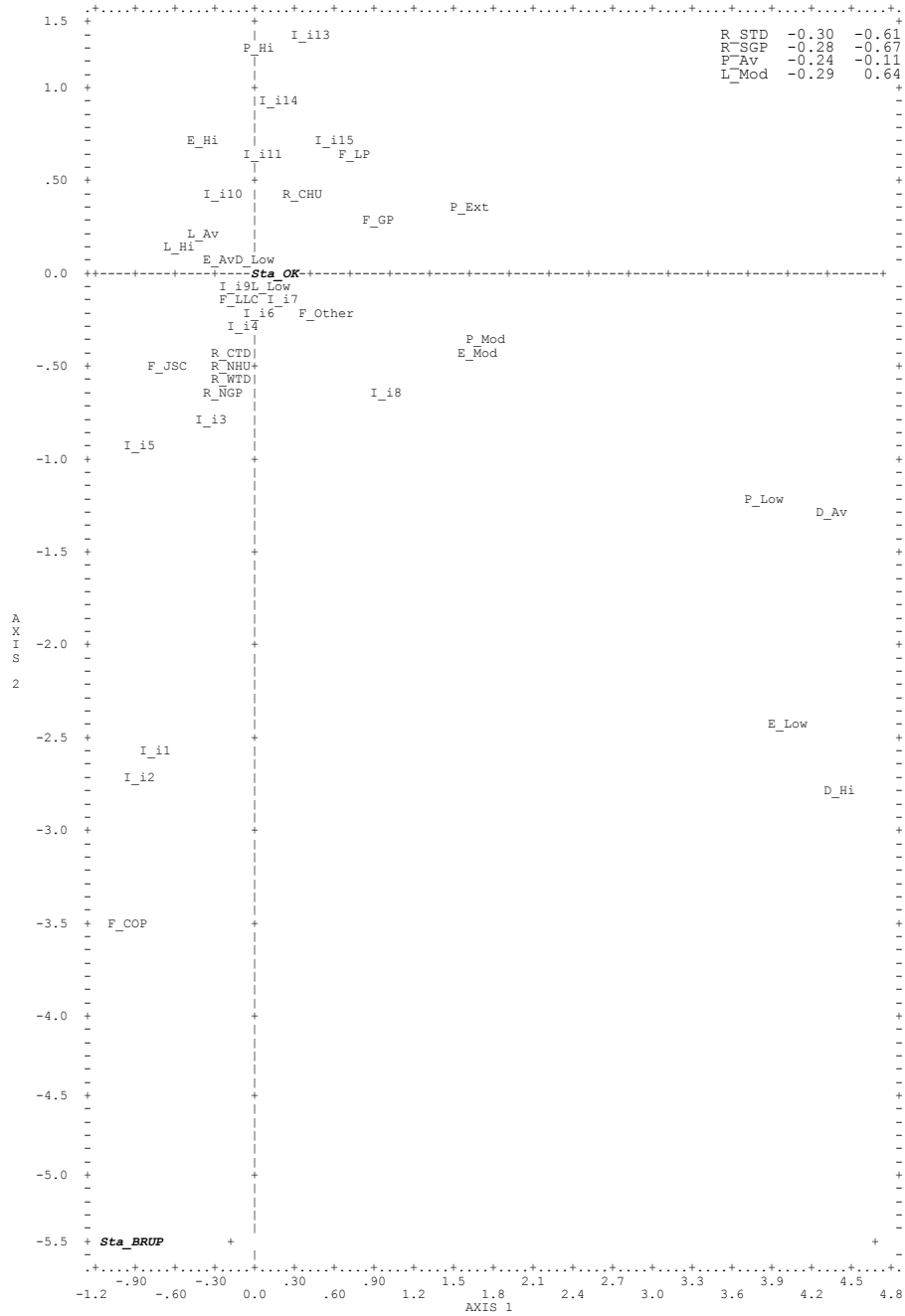
The total inertia considering the indicator matrix with 45 columns is $(45/8-1)=4.625$. In this case $1/Q=1/8=0.125$ hence the CA axes with squared singular values greater than 0.125 are worth being extracted. The percentages of the total inertia accounted for by the three leading axes are: $\mu_1^2=0.248$ (5.4%), $\mu_2^2=0.184$ (4%), $\mu_3^2=0.159$ (3.4%). The number of meaningful axes actually extracted is 3, no matter the cumulative percentage of inertia accounted for by them. Table 8 provides information given in the column coordinates. In the table NAME identifies the column concerned and MASS stands for the category total as a proportion of all cases. Again, according to the definitions introduced earlier QLT is a quality-measure of how well the distance from the origin of this point in the reduced dimension (three in this case) represents the full distance from the origin. The contribution of the given category to the total inertia is measured by INR. Furthermore, attributes for each extracted dimension are as follows:

- FACTOR: the category coordinate or column score for the corresponding axis.
- COR2: the 'squared correlation' indicates how well the distance of the point along that axis from the origin represents the total distance of the point from the origin. The sum of COR2 values for the axes extracted equals QLT.
- CTR shows the category's relative contribution to the inertia accounted for by that axis.

When column profiles in the indicator matrix are similar, the corresponding category points in the graphical display will tend to be close together. Thus, the same or similar firms will be in categories represented by adjacent points. In addition, Figure 1 labels category areas using the FACTOR values on Axis 1 and Axis 2, Figure 2 shows the category positions in the plain of Axis 1 and Axis 3 and, finally Figure 3 plots the column points along Axis 2 and Axis 3. The category point could also be interpreted as the category mean scores for the first three axes. Thus, the interpretation of the axes is as follows.

- The first axis contrasts: P_Low, E_Low, D_High, I_i8, F_GP, F_LP, P_Mod, P_Ext with positive coordinates against Sta_BRUPT, F_COP, F_JSC, I_i1, I_i2, I_i5, L_Hi, with negative coordinates.
- The second axis contrasts: I_i13, P_High, I_i14, I_i15, E_High, F_LP, I_i11 with positive coordinates against Sta_BRUPT, F_COP, D_High, E_Low, I_i1, I_i2, P_LOW, I_i5, D_Av with negative coordinates.
- The third axis contrasts: Sta_BRUPT, F_COP, I_i1, I_i2, E_High, with positive coordinates against D_High, E_Low, I_i5, I_i3 with negative coordinates.

Figure 1. Correspondences between the column categories on Axis 1, 2
 $X=5.36\%$, $Y=3.98\%$



Note: On Figures 1,2,3,4,5,6 the labels and coordinates of points which would overwrite points already plotted are given in the right hand upper corner of the display.

Figure 2. Correspondences between the column categories on Axis 1, 3
 $X=5.36\%$, $Y=3.44\%$

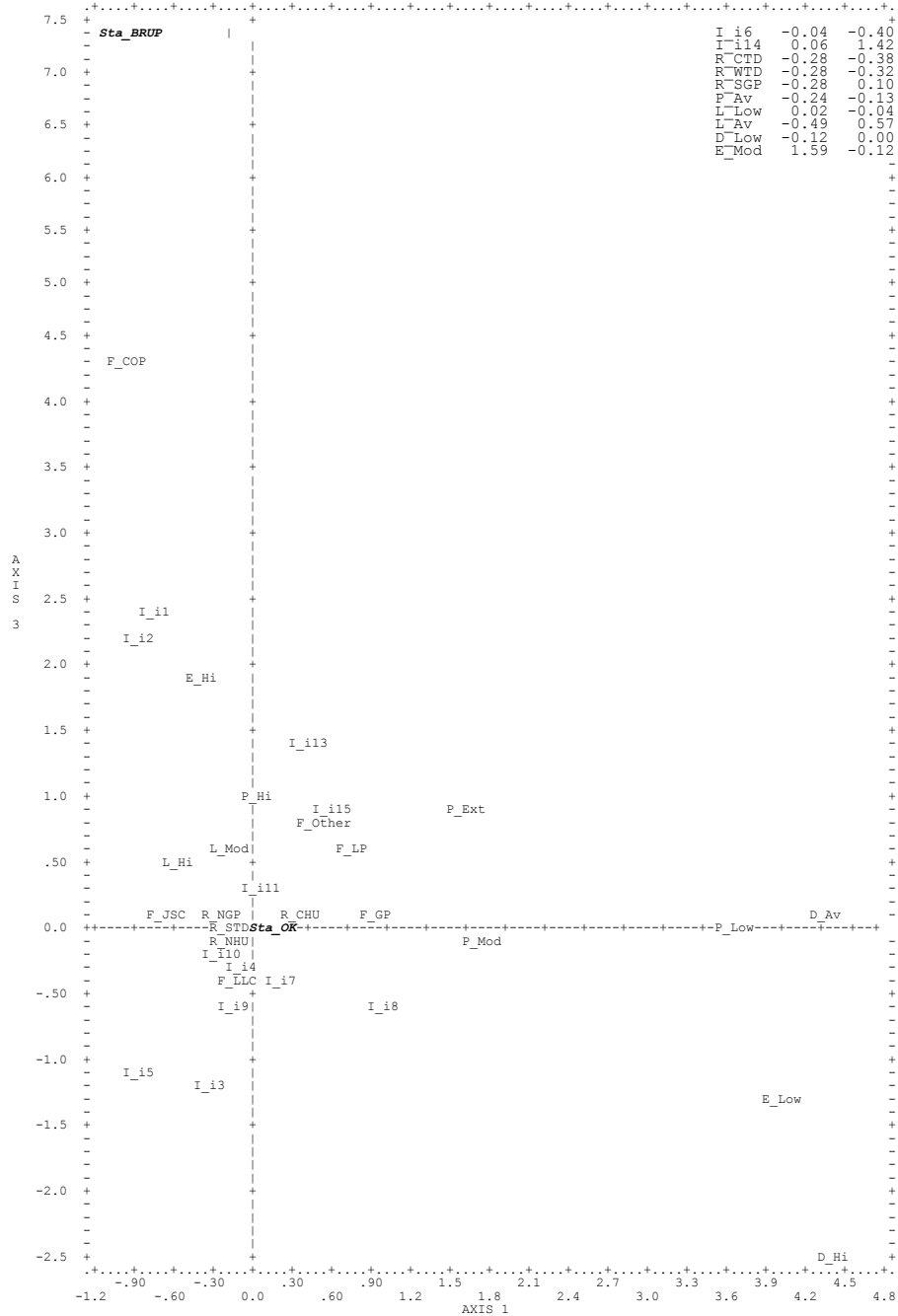


Table 8

Report on the categories as column points of the indicator matrix

NAME	MASS	QLT	INR	FACTOR	COR2	CTR	FACTOR	COR2	CTR	FACTOR	COR2	CTR
				AXIS 1			AXIS 2			AXIS 3		
Sta_OK	0.125	0.164	0.000	0.002	0.003	0.000	0.010	0.057	0.000	-0.014	0.105	0.000
Sta_BRUP	0.000	0.164	0.125	-1.158	0.003	0.001	-5.471	0.057	0.038	7.445	0.105	0.082
F_Other	0.002	0.014	0.123	0.331	0.002	0.001	-0.220	0.001	0.001	0.837	0.011	0.009
F_COP	0.002	0.545	0.123	-1.093	0.020	0.010	-3.496	0.208	0.139	4.321	0.317	0.245
F_GP	0.001	0.007	0.124	0.842	0.006	0.003	0.270	0.001	0.000	0.116	0.000	0.000
F_LP	0.034	0.462	0.091	0.658	0.165	0.060	0.645	0.158	0.078	0.603	0.138	0.079
F_LLC	0.084	0.475	0.041	-0.248	0.125	0.021	-0.168	0.057	0.013	-0.380	0.293	0.076
F_JSC	0.002	0.011	0.123	-0.764	0.008	0.004	-0.481	0.003	0.002	0.135	0.000	0.000
I_i1	0.005	0.508	0.120	-0.843	0.028	0.013	-2.564	0.257	0.168	2.390	0.223	0.169
I_i2	0.000	0.008	0.125	-0.943	0.001	0.000	-2.721	0.005	0.003	2.159	0.003	0.002
I_i3	0.000	0.003	0.125	-0.409	0.000	0.000	-0.792	0.001	0.001	-1.183	0.002	0.002
I_i4	0.019	0.030	0.106	-0.201	0.007	0.003	-0.253	0.011	0.007	-0.253	0.011	0.008
I_i5	0.000	0.008	0.125	-0.969	0.002	0.001	-0.938	0.002	0.002	-1.062	0.003	0.002
I_i6	0.010	0.017	0.115	-0.039	0.000	0.000	-0.179	0.003	0.002	-0.403	0.014	0.010
I_i7	0.039	0.109	0.086	0.110	0.006	0.002	-0.178	0.015	0.007	-0.442	0.089	0.048
I_i8	0.005	0.059	0.120	0.879	0.030	0.015	-0.645	0.016	0.011	-0.590	0.013	0.010
I_i9	0.005	0.017	0.120	-0.258	0.003	0.001	-0.098	0.000	0.000	-0.580	0.014	0.010
I_i10	0.000	0.001	0.125	-0.371	0.000	0.000	0.400	0.000	0.000	-0.185	0.000	0.000
I_i11	0.032	0.178	0.093	-0.055	0.001	0.000	0.657	0.146	0.074	0.305	0.031	0.018
I_i13	0.001	0.043	0.124	0.283	0.001	0.000	1.286	0.019	0.013	1.439	0.024	0.018
I_i14	0.003	0.071	0.122	0.063	0.000	0.000	0.964	0.022	0.015	1.419	0.048	0.037
I_i15	0.006	0.073	0.119	0.452	0.010	0.005	0.721	0.026	0.017	0.860	0.037	0.028
R_CHU	0.070	0.344	0.055	0.243	0.074	0.017	0.456	0.262	0.079	0.075	0.007	0.002
R_CTD	0.009	0.029	0.116	-0.282	0.006	0.003	-0.395	0.012	0.007	-0.382	0.011	0.008
R_WTD	0.009	0.042	0.116	-0.279	0.006	0.003	-0.595	0.028	0.017	-0.321	0.008	0.006
R_STD	0.008	0.033	0.117	-0.300	0.007	0.003	-0.606	0.027	0.017	0.019	0.000	0.000
R_NHU	0.008	0.024	0.117	-0.302	0.006	0.003	-0.508	0.017	0.011	-0.126	0.001	0.001
R_NGP	0.010	0.049	0.115	-0.387	0.013	0.006	-0.632	0.035	0.022	0.070	0.000	0.000
R_SGP	0.011	0.053	0.114	-0.280	0.008	0.004	-0.666	0.044	0.027	0.100	0.001	0.001
P_Low	0.003	0.426	0.122	3.697	0.383	0.188	-1.243	0.043	0.029	0.011	0.000	0.000
P_Mod	0.007	0.155	0.118	1.612	0.148	0.071	-0.327	0.006	0.004	-0.150	0.001	0.001
P_Av	0.100	0.349	0.025	-0.237	0.226	0.023	-0.111	0.050	0.007	-0.134	0.073	0.011
P_Hi	0.014	0.316	0.111	-0.057	0.000	0.000	1.240	0.194	0.117	0.985	0.122	0.085
P_Ext	0.001	0.018	0.124	1.478	0.013	0.006	0.322	0.001	0.000	0.891	0.005	0.004
L_Low	0.118	0.050	0.007	0.018	0.005	0.000	-0.037	0.023	0.001	-0.036	0.022	0.001
L_Mod	0.007	0.048	0.118	-0.291	0.005	0.002	0.638	0.023	0.015	0.606	0.021	0.015
L_Av	0.000	0.001	0.125	-0.493	0.000	0.000	0.237	0.000	0.000	0.574	0.001	0.000
L_Hi	0.000	0.001	0.125	-0.681	0.000	0.000	0.139	0.000	0.000	0.491	0.000	0.000
D_Low	0.122	0.546	0.003	-0.118	0.500	0.007	0.036	0.046	0.001	-0.002	0.000	0.000
D_Av	0.003	0.542	0.122	4.237	0.497	0.244	-1.284	0.046	0.030	0.094	0.000	0.000
D_Hi	0.000	0.005	0.125	4.337	0.003	0.001	-2.770	0.001	0.001	-2.512	0.001	0.001
E_Low	0.000	0.016	0.125	3.928	0.010	0.005	-2.439	0.004	0.003	-1.314	0.001	0.001
E_Mod	0.022	0.575	0.103	1.587	0.537	0.223	-0.407	0.035	0.020	-0.118	0.003	0.002
E_Av	0.103	0.577	0.022	-0.342	0.539	0.048	0.088	0.036	0.004	0.023	0.002	0.000
E_Hi	0.000	0.006	0.125	-0.456	0.000	0.000	0.715	0.001	0.001	1.871	0.005	0.004

Note: Measures with relatively high values are emphasized in the table.

As the number of the OK firms dominates the frequency of the BRUPT firms the average profile of the OK group is almost the same as the centroid of all firms. Hence, on the plots the OK category is well represented by the origin. Thus, categories located around the origin (but not too close to it because of the reconstitution formula /18/) can be associated with no bankruptcy. Categories such as F_GP, P_Ext, F_LP, I_i15 are obviously far from bankruptcy in the plane of Axis 1 and Axis 2 while category F_COP and I_i2 are obviously close to Sta_BRUPT. On the other hand, category BRUPT moves far from the origin and pulls some predictor categories in the same direction on the respective axes such as I_i1, I_i3, I_i5, F_JSC on the first axis and D_High, E_Low, P_Low, D_Av on the second axis. Considering the transition formula /15/ and recalling its interpretation the distances among the points on the plots show how strongly or slightly are these points associated with the category BRUPT.

PREDICTIVE MAPS

It is supported by Multiple Correspondence Analysis to add further points to the plots representing the categories of additional or supplementary variables. Supplementary variables correspond to dependent or response variables. Simply we plot the column points of the indicator matrix in a two dimensional space to explore dependencies among the predictor categories. This is called a *predictive map*. Then we project the supplementary row profile onto this predictive map (using the transition formula /15/). The positions of the supplementary categories in relation to the positions of the predictor categories reveal their dependencies or independencies. In our case the OK and BRUPT categories play the role of supplementary categories desired to be discriminated on the predictive map while the position of the PROC supplementary category (an average of the individual PROC-member points) is compared with the positions of OK and BRUPT. Obviously, all PROC members are excluded from the computations of the predictive map. Nevertheless, an additional individual firm could also be projected onto the predictive map letting us investigate its neighbours whether they are mostly bankrupt or not. However, this latter concept of visual classification would need a ‘scatter plot of firms’ with the difficulty in its interpretation due to the large number of the firms to be plotted actually in the display.

Table 9

Report on the predictor categories as column points of the indicator matrix

NAME	MASS	QLT	INR	FACTOR	COR2	CTR	FACTOR	COR2	CTR	FACTOR	COR2	CTR
				AXIS	AXIS	1	AXIS	2	AXIS	3		
F_Other	0.002	0.019	0.141	0.333	0.002	0.001	-0.183	0.001	0.000	1.020	0.017	0.013
F_COP	0.002	0.532	0.140	-1.029	0.018	0.009	-3.019	0.155	0.105	4.599	0.359	0.284
F_GP	0.001	0.007	0.142	0.840	0.006	0.003	0.293	0.001	0.001	0.051	0.000	0.000
F_LP	0.039	0.464	0.103	0.656	0.164	0.060	0.704	0.189	0.094	0.541	0.111	0.065
F_LLC	0.096	0.478	0.047	-0.249	0.126	0.021	-0.206	0.086	0.020	-0.362	0.266	0.071
F_JSC	0.002	0.010	0.141	-0.753	0.007	0.004	-0.414	0.002	0.002	-0.006	0.000	0.000
I_i1	0.005	0.591	0.137	-0.817	0.026	0.013	-2.422	0.229	0.152	2.929	0.335	0.259
I_i2	0.000	0.006	0.143	-0.890	0.000	0.000	-2.344	0.003	0.002	1.896	0.002	0.002
I_i3	0.000	0.004	0.143	-0.402	0.000	0.000	-0.837	0.001	0.001	-1.327	0.003	0.002
I_i4	0.022	0.032	0.121	-0.199	0.007	0.003	-0.267	0.013	0.007	-0.267	0.013	0.009
I_i5	0.000	0.007	0.142	-0.971	0.002	0.001	-1.084	0.003	0.002	-0.837	0.002	0.001
I_i6	0.011	0.017	0.131	-0.039	0.000	0.000	-0.216	0.004	0.003	-0.386	0.013	0.010
I_i7	0.045	0.112	0.098	0.109	0.005	0.002	-0.217	0.022	0.010	-0.430	0.085	0.047
I_i8	0.005	0.059	0.138	0.880	0.030	0.015	-0.697	0.019	0.012	-0.513	0.010	0.008
I_i9	0.006	0.017	0.137	-0.260	0.003	0.001	-0.154	0.001	0.001	-0.567	0.013	0.010
I_i10	0.000	0.001	0.143	-0.375	0.000	0.000	0.376	0.000	0.000	-0.207	0.000	0.000
I_i11	0.036	0.178	0.107	-0.058	0.001	0.000	0.695	0.163	0.084	0.202	0.014	0.008
I_i13	0.002	0.046	0.141	0.278	0.001	0.000	1.399	0.022	0.015	1.409	0.023	0.018
I_i14	0.003	0.080	0.140	0.060	0.000	0.000	1.057	0.027	0.018	1.489	0.053	0.042
I_i15	0.007	0.075	0.136	0.450	0.010	0.005	0.795	0.032	0.021	0.816	0.033	0.025
R_CHU	0.080	0.360	0.063	0.241	0.073	0.016	0.477	0.286	0.087	-0.012	0.000	0.000
R_CTD	0.010	0.029	0.133	-0.281	0.006	0.003	-0.432	0.014	0.009	-0.351	0.009	0.007
R_WTD	0.010	0.042	0.132	-0.277	0.006	0.003	-0.640	0.032	0.020	-0.220	0.004	0.003
R STD	0.010	0.038	0.133	-0.298	0.006	0.003	-0.635	0.029	0.019	0.180	0.002	0.002
R_NHU	0.009	0.024	0.134	-0.299	0.006	0.003	-0.527	0.018	0.012	-0.056	0.000	0.000
R_NGP	0.012	0.056	0.131	-0.385	0.013	0.006	-0.658	0.038	0.024	0.241	0.005	0.004
R_SGP	0.013	0.057	0.130	-0.275	0.008	0.003	-0.671	0.045	0.028	0.212	0.004	0.003
P_Low	0.004	0.427	0.139	3.706	0.385	0.189	-1.223	0.042	0.028	0.172	0.001	0.001
P_Mod	0.008	0.156	0.135	1.615	0.148	0.071	-0.313	0.006	0.004	-0.182	0.002	0.001
P_Av	0.114	0.349	0.028	-0.237	0.226	0.023	-0.126	0.064	0.009	-0.122	0.060	0.010
P_Hi	0.016	0.319	0.127	-0.062	0.000	0.000	1.328	0.222	0.136	0.875	0.097	0.069
P_Ext	0.001	0.018	0.142	1.484	0.013	0.007	0.446	0.001	0.001	0.779	0.004	0.003
L_Low	0.135	0.052	0.008	0.018	0.006	0.000	-0.039	0.026	0.001	-0.035	0.021	0.001
L_Mod	0.008	0.051	0.135	-0.294	0.005	0.002	0.679	0.026	0.017	0.600	0.020	0.015
L_Av	0.000	0.001	0.143	-0.496	0.000	0.000	0.252	0.000	0.000	0.654	0.001	0.001
L_Hi	0.000	0.000	0.143	-0.675	0.000	0.000	0.231	0.000	0.000	0.282	0.000	0.000
D_Low	0.139	0.547	0.004	-0.118	0.502	0.007	0.035	0.044	0.001	-0.006	0.001	0.000
D_Av	0.004	0.544	0.139	4.247	0.499	0.245	-1.245	0.043	0.029	0.235	0.002	0.001
D_Hi	0.000	0.005	0.143	4.347	0.003	0.001	-2.973	0.001	0.001	-1.971	0.001	0.000
E_Low	0.000	0.015	0.143	3.939	0.010	0.005	-2.556	0.004	0.003	-0.870	0.001	0.000
E_Mod	0.025	0.575	0.118	1.590	0.539	0.224	-0.400	0.034	0.019	-0.100	0.002	0.001
E_Av	0.117	0.577	0.025	-0.342	0.542	0.049	0.086	0.034	0.004	0.019	0.002	0.000
E_Hi	0.000	0.006	0.143	-0.445	0.000	0.000	0.944	0.001	0.001	1.654	0.004	0.003

Note: Measures with relatively high values are emphasized in the table.

The total inertia considering the indicator matrix with only the 43 columns is $(43/7-1)=5.1429$. The percentages of the total inertia accounted for by the 3 leading axes are: $\mu_1^2=0.283$ (5.5%), $\mu_2^2=0.208$ (4.0%), $\mu_3^2=0.178$ (3.5%). The ‘quality’ measures and the coordinates for the predictive maps are given in Table 9 and for the supplementary points in Table 10. The 0.291 PROC coordinate on Axis 1 for example, is the weighted average of the Axis 1 coordinates given in Table 9 using the PROC average row profile values as a weighting scheme. Based on this interpretation and the fairly high QLT values ‘OK firms’ are still represented by the origin, while ‘BRUPT firms’ mostly come from the categories close to $-0.283, 0.745, 0.688$ on the respective axes and ‘PROC firms’ belong to the categories with coordinates near $0.291, 0.411, -0.102$.

Table 10

Supplementary profiles (centroids) of groups OK, BRUPT, PROC

NAME	QLT	FACTOR AXIS	COR2 1	FACTOR AXIS	COR2 2	FACTOR AXIS	COR2 3
OK	0.721	0.001	0.050	0.001	0.321	-0.001	0.350
BRUPT	0.721	-0.296	0.050	-0.748	0.321	0.781	0.350
PROC	0.740	0.291	0.233	-0.426	0.499	-0.054	0.008

Thus, the predictive categories associated with category OK are: F_LLC, I_i4, I_i6, I_i7, I_9, I_i10, R_CHU, R_CTD, L_Low, D_Low, E_Av. Further, the predictive categories associated with category BRUPT based on Axis1 and Axis 2 are: I_i2, R_WTD, R_STD, R_NHU, R_NGP, R_SGP, and based on Axis1 and Axis 3 are: L_Mod, L_Av.

As mentioned earlier, a single firm can also be classified by its projection on the predictive map. Apparently, because we have 7 predictor variables any firm’s row profile (as its weighting scheme) contains a value of $1/7$ at each column position that this firm belongs to and a zero anywhere else. Hence, only the column coordinates associated with the firm concerned play a role in an individual’s prediction. Once the firm’s coordinates are given, we can compare them with the centroid of any supplementary category.

For example the first coordinate of a firm with F=LLC, I=i1, R=CHU, P=Low, L=Low, D=High, E=Low profile is calculated as the simple average of FACTOR_1 the second coordinate is the simple average of FACTOR_2 and the third coordinate is the simple average of FACTOR_3 averaged on the corresponding categories only as follows:

$$\mu_1 \cdot \text{FACTOR}_1 = (-0.249 - 0.817 + 0.241 + 3.706 + 0.018 + 4.347 + 3.939) / 7 = 1.5978$$

$$\mu_2 \cdot \text{FACTOR}_2 = (-0.206 - 2.422 + 0.477 - 1.223 - 0.039 - 2.973 - 2.556) / 7 = -1.2774$$

$$\mu_3 \cdot \text{FACTOR}_3 = (-0.362 + 2.929 - 0.012 + 0.172 - 0.035 - 1.971 - 0.870) / 7 = -0.0213$$

and the standardized coordinates are:

$$\text{FACTOR}_1 = 1.5978 / 0.283^{1/2} = 3.003$$

$$\text{FACTOR}_2 = -1.2774 / 0.208^{1/2} = -2.801.$$

$$\text{FACTOR}_3 = -0.0213 / 0.178^{1/2} = -0.050.$$

Figure 4. Map of the predictive variables by Axes 1,2
 X=5.50 %, Y=4.04 %

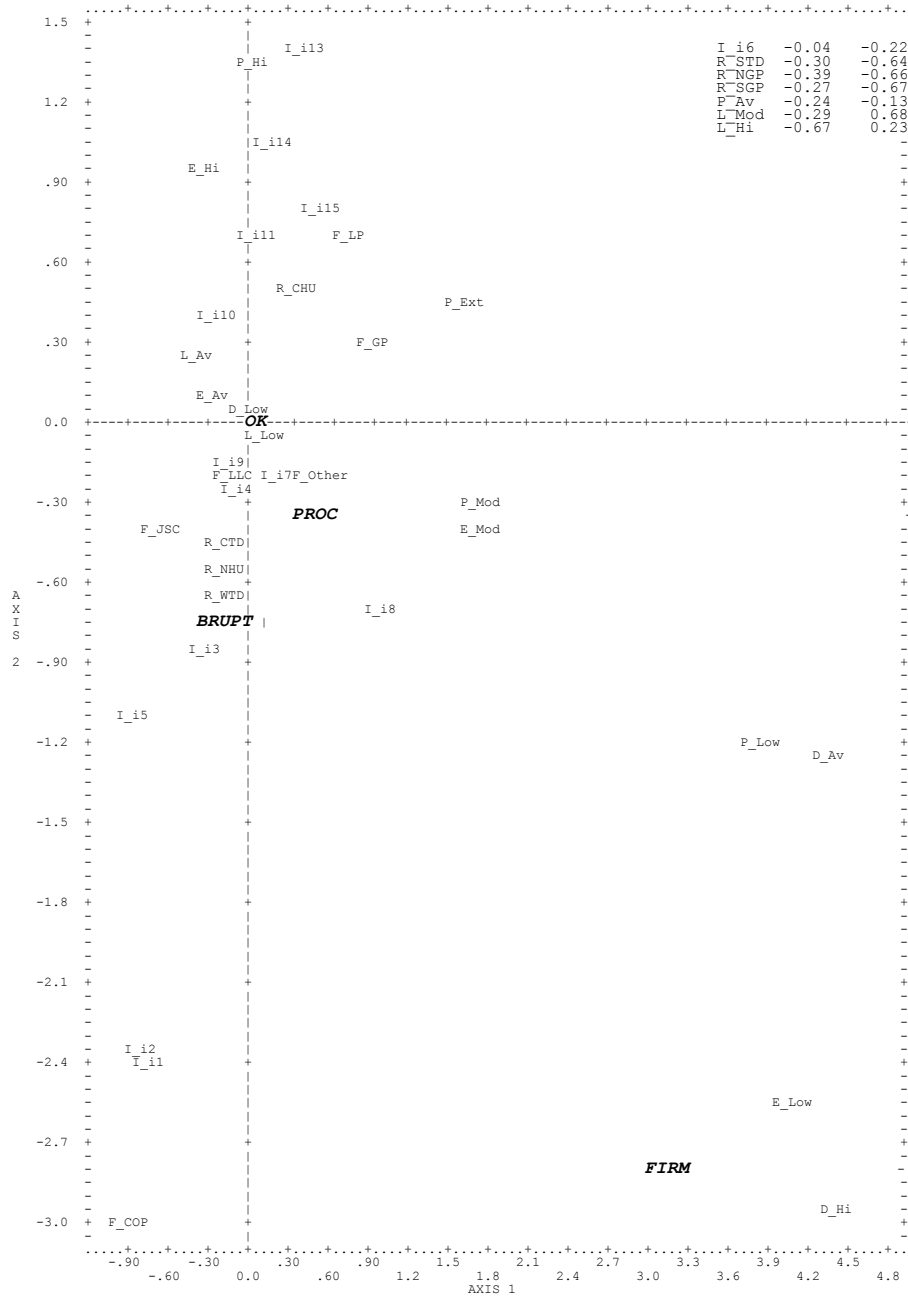


Figure 5. Map of the predictive variables by Axes 1, 3
 X=5.50 %, Y=3.46 %

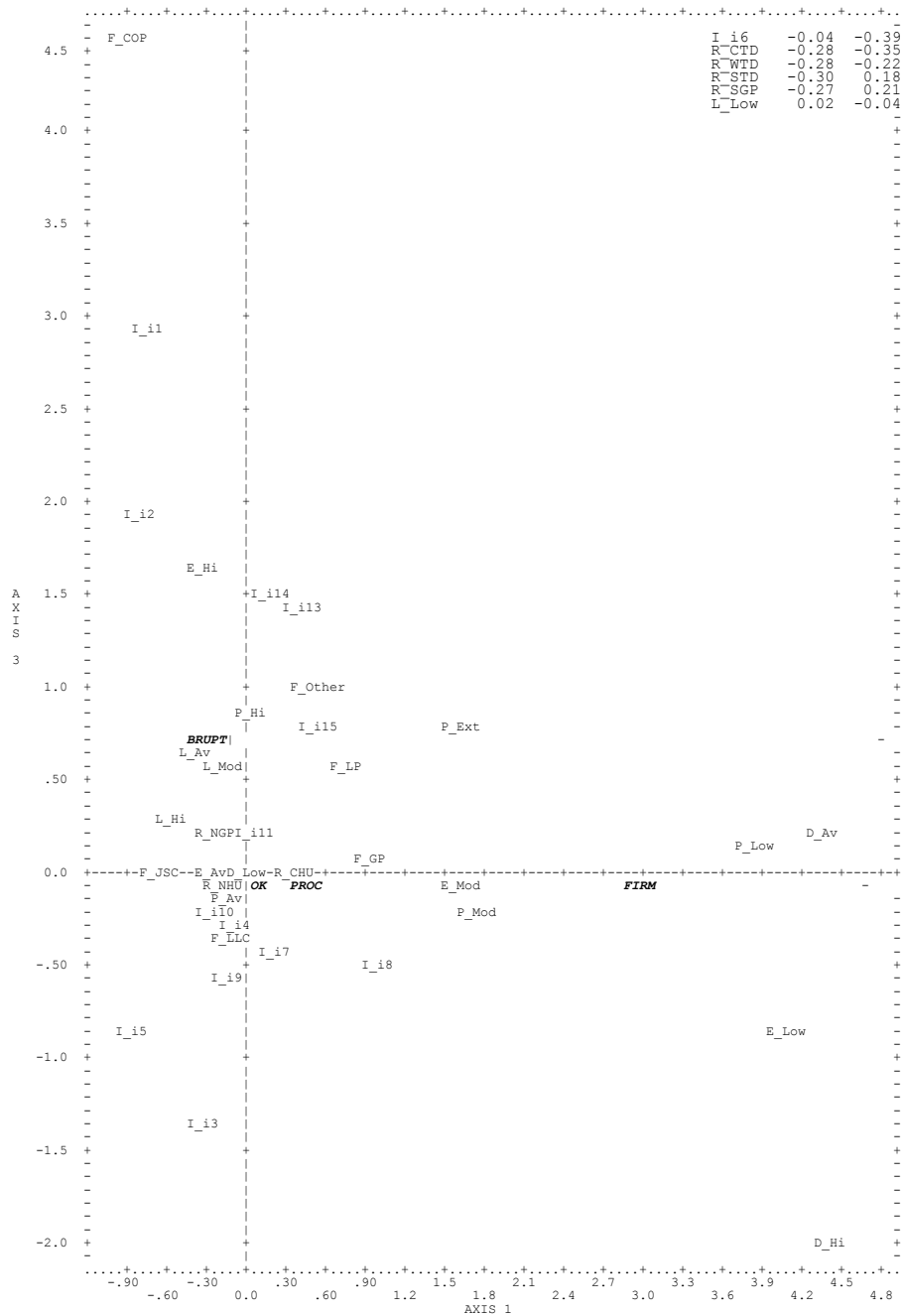
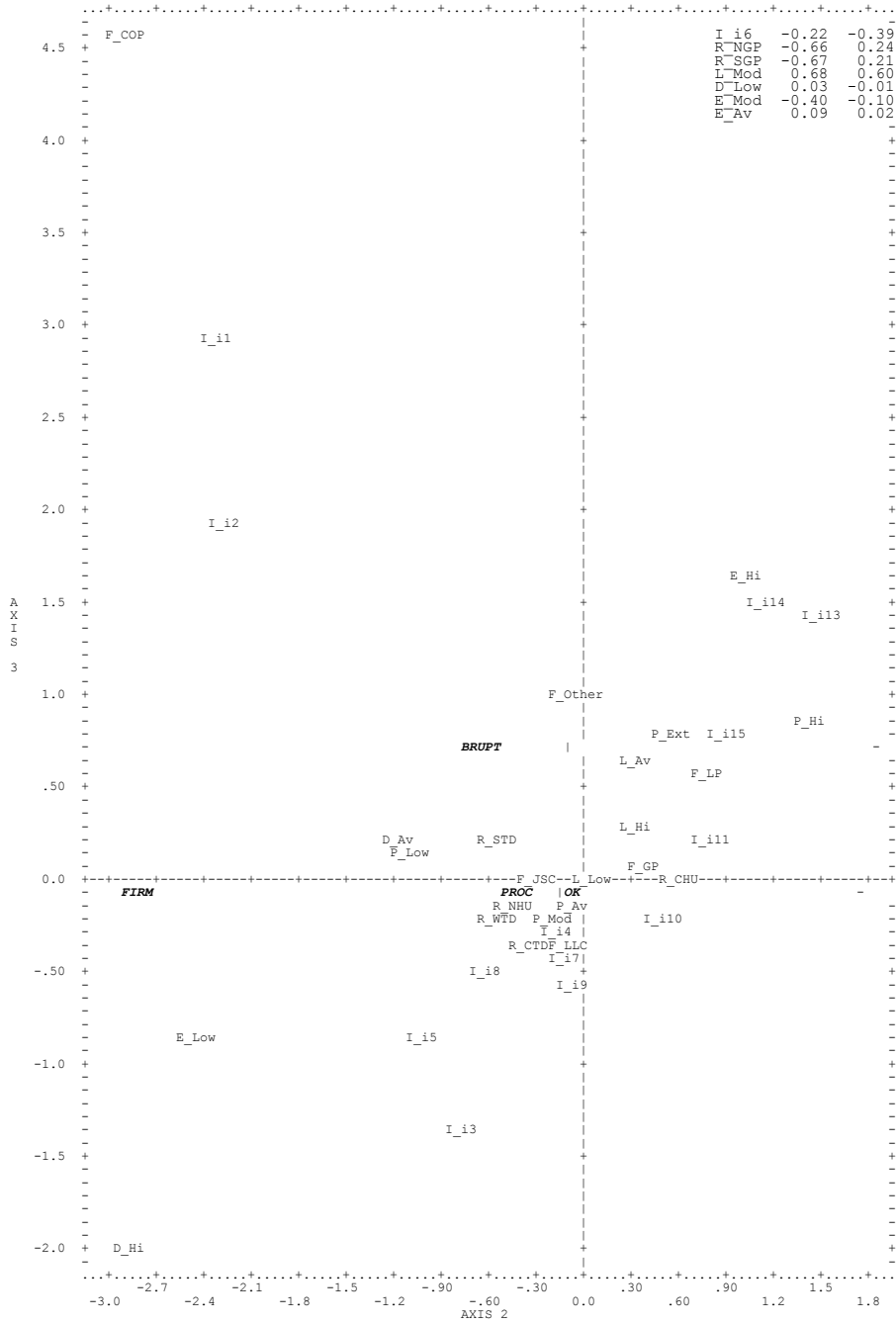


Figure 6. Map of the predictive variables by Axes 2,3
 X=4.04 %, Y=3.46 %



It should be noted, that L=Low (the lowest category of liquidity) could be omitted from the calculation leaving the factor scores calculated almost unchanged. Plotting this profile on Figure 4, 5, 6 serves for a classification of the firm of interest. According to these figures a clear OK, PROC, BRUPT line can be explored and a limited liability company from Central Hungary and from industry 1, with low profitability, low liquidity, high debt ratio and low equity ratio (labelled FIRM) exhibits extreme coordinates in the direction of those who are currently bankrupt.

REFERENCES

- DIXON, W. J. (chief editor) (1992): *BMDP statistical software manual*. Vol. 2, University Press of California, Berkely, Los Angeles, Oxford.
- GREENACRE, M. J. (1984): *Theory and application of correspondence analysis*. Academic Press, Inc., London.
- HIGGS, N. T. (1990): Practical and innovative uses of correspondence analysis. *The Statistician*, Vol. 40. p. 183–194.