

## ROW AND COLUMN MATRICES IN MULTIPLE CORRESPONDENCE ANALYSIS WITH ORDERED CATEGORICAL AND DICHOTOMOUS VARIABLES

### Article history

Received

11 February 2015

Received in revised form

25 March 2015

Accepted

15 January 2016

Thanoon Y. Thanoon<sup>a,b\*</sup>, Robiah Adnan<sup>a</sup><sup>a</sup>Faculty of Science, Universiti Teknologi Malaysia, Malaysia<sup>b</sup>Northern Technical University, Technical College of Management, Mosul, Iraq

\*Corresponding author

ytthaon2@live.utm.my

### Abstract

In multiple correspondence analysis, whenever the number of variables exceeds the number of observations, row matrix should be used, but if the number of variables is less than the number of observations column matrix is the suitable procedure to follow. One of the following matrices (rows, columns) leads to loss of information that can be found by the other method, therefore, this paper developed a proposal to overcome this problem, which is: to find a shortcut method allowing the use of the results of one matrix to obtain the results of the other matrix. Taking advantage of all information available, the phenomenon was studied. Some of these results are: Eigenvectors, factor loadings and factor scores based on ordered categorical and dichotomous data. This method is illustrated by using a real data set. Results were obtained by using Minitab program. As a result, it is possible to shortcut transformation between the results of row and column matrices depending on factor loadings and factor scores of the row and column matrices.

Keywords: Multiple correspondence analysis, row matrix, column matrix, ordered categorical data, dichotomous data

### Abstrak

Dalam Analisis Korespondensi Berganda, matriks baris (row) boleh digunakan jika bilangan pemboleh ubah melebihi bilangan pemerhatian. Sebaliknya, jika bilangan pemboleh ubah adalah kurang daripada bilangan pemerhatian maka matriks lajur (column) adalah lebih sesuai digunakan. Walau bagaimanapun, penggunaan salah satu daripada matriks berikut (baris, lajur) menjurus kepada kehilangan maklumat yang boleh dicapai melalui kaedah capaian yang lain. Justeru itu, tujuan pembangunan kertas kerja cadangan ini adalah untuk mencari kaedah jalan pintas (shortcut) yang membenarkan penggunaan keputusan satu matriks untuk mendapatkan keputusan matriks yang lain. Fenomenon yang dikaji mengambil kira semua maklumat yang ada, dan hasil keputusan daripada sebahagian kajian tersebut adalah vektor eigen (eigenvectors), faktor beban dan faktor skor berdasarkan data berkategori tertib, dan data dikotomi. Kaedah ini digambarkan dengan penggunaan set data sebenar, dan program Minitab digunakan untuk mendapatkan keputusan yang dikehendaki. Hasil keputusan kajian ini menunjukkan bahawa terdapat kemungkinan untuk pintasan transformasi di antara keputusan daripada baris, dan lajur matriks yang bergantung kepada faktor beban dan faktor skor baris dan lajur matriks.

Kata kunci: Analisis Korespondensi Berganda, Matriks Baris, Matriks Lajur, data berkategori tertib, data dikotomi

© 2016 Penerbit UTM Press. All rights reserved

## 1.0 INTRODUCTION

Factor analysis is designed for interval or ratio data; that is, measurements made on a continuous numerical scale and are not appropriate for enumerative data, such as the numbers of fossils, and

sometimes the number of fractures. Nominal or ordinal observations may be all that are available, and in some instances it seems desirable to process these using an Eigenvalue technique which is similar to factor analysis [1], [2].

Count data are available and very common in social sciences. Surveys by questionnaire, for example, result in responses that can only be tallied in categories. Consequently, most of the research on the use of Eigenvalue methods for analyzing such data has been done by sociologists and statisticians working on sociological problems. These data may be summarized conveniently in the form of contingency tables; the first analysis of the type we are considering was known as "contingency table analysis". More recently, Benzécri and Benzécri [3] have written extensively on the subject, and Benzécri's term "correspondence analysis" has come to be widely used. His work provides the basis for many applications in geology. In these geological applications, the methods of Benzécri and his predecessors have been extensively modified. Hill [4] discusses the history of correspondence analysis, and the interrelationships between the works of various authors. A detailed discussion of correspondence analysis and its extensions is contained in the monograph by Lebart *et al.*, [5].

Useful treatment of correspondence analysis has been given by Greenacre [6], Khattree and Naik [7], Gower and Hand [8], and Benzécri [9]. More statistically sound methods for correspondence analysis have been proposed by Heiser [10] and Greenacre [6].

The paper is organized as follows. Correspondence Analysis is described in section 2. Multiple correspondence analysis is explained in Section 3. Rows and columns profiles are discussed in section 4. Shortcut Transformation between column matrix (C) and row matrix (R) matrices are presented in section 5. Real example to illustrate the method used is presented in Section 6. Statistical analysis and empirical results are discussed in section 7. Some concluding remarks are given in section 8.

## 2.0 CORRESPONDENCE ANALYSIS

Correspondence analysis provides tools for analyzing the associations between rows and columns of contingency tables. A contingency table is a two-entry frequency table where the joint frequencies of two qualitative variables are reported. For instance a 2x2 table could be formed by observing from a sample of  $n$  individuals two qualitative variables: the individual's gender and whether the individual smokes. The table reports the observed joint frequencies. In general  $(n \times p)$  tables where  $n$  is the number of observation and  $p$  is the number of variables which may be considered [11], [12].

To test for significance of association of the two categorical variables in a contingency table, we could use a chi-square test which is representing an asymptotic approach. Since correspondence analysis is associated with the chi-square approach, if a contingency table has some cell frequencies that are small or zero, the chi-square approximation is not very

satisfactory. In this case, some categories can be combined to increase the cell frequencies.

Correspondence Analysis may be useful in identifying the categories that are similar, which we may thereby wish to combine. In correspondence analysis, we plot a point for each row and a point for each column of the contingency table. These points are, in effect, projections of the rows and columns of the contingency table onto a two dimensional Euclidean space. The goal is to preserve as far as possible the relationship of the rows or (columns) to each other in a two-dimensional space. If two row points are close together, the profiles of the two rows (across the columns) are similar. Likewise, two column points that are close together represent columns with similar profiles across the rows. If a row point is close to a column point, this combination of categories of the two variables occurs more frequently than would occur by chance if the two variables were independent.

The idea is to extract the indices in decreasing order of importance so that the main information of the table can be summarized in spaces with smaller dimensions. For instance, if only two factors (indices) are used, the results can be shown in two-dimensional graphs, showing the relationship between the rows and the columns of the table [13].

## 3.0 MULTIPLE CORRESPONDENCE ANALYSIS

Correspondence Analysis of a two-way contingency table can be extended to a three way or higher-order multi-way table, by the method of multiple correspondence analysis. We will obtain a two-dimensional graphical display of the information in the multi-way contingency table [14].

Many articles have mentioned Multiple Correspondence Analysis including (Abdi & Valentin, [15]; Ambrogi *et al.* [16]; Loslever *et al.*, [17]; Du [14]; Greenacre & Blasius, [18]; Le Roux & Rouanet, [19]; Panagiotakos & Pitsavos, [20], and Reis *et al.*, [21]).

## 4.0 ROWS AND COLUMNS PROFILES

A contingency table with  $a$  rows and  $b$  columns is represented in Table 1. The entries  $n_{ij}$  are the counts or frequencies for every two-way combination of row and column (every cell). The marginal totals are shown using the familiar dot notation [12].

$$n_{i.} = \sum_{j=1}^b n_{ij} \text{ and } n_{.j} = \sum_{i=1}^a n_{ij} . \text{ The overall total}$$

frequency is denoted by  $n$  instead of  $n_{i.}$  and  $n_{.j}$  For simplicity:

$$n = \sum_{ij} n_{ij}$$

The frequencies  $n_{ij}$  in a contingency table can be converted to relative frequencies  $p_{ij}$  by dividing by  $n$ :  $p_{ij} = n_{ij} / n$ . The matrix of relative frequencies is called

the correspondence matrix and is denoted by  $P$  as in the Tables (1) and (2), respectively:

$$P = [P_{ij}] = [n_{ij}/n] \tag{1}$$

In Table 2 we show the contingency table in Table 1 converted to a correspondence matrix.

The last column of Table 2 contains the row sums

$$p_i = \sum_{j=1}^b p_{ij} .$$

This column vector is denoted by  $r$  and

can be obtained as

$$r = p_j = (p_1, p_2, \dots, p_a)' = (n_{.1}/n, n_{.2}/n, \dots, n_{.a}/n)'$$

where  $j$  is  $a \times 1$  vector. Similarly, the last row of Table 2

contains the column sums  $p_{.j} = \sum_{i=1}^a p_{ij}$ . This row

vector is denoted by  $c'$  and can be denoted as:

$$c' = j'p = (p_{.1}, p_{.2}, \dots, p_{.a})' = (n_{.1}/n, n_{.2}/n, \dots, n_{.a}/n)'$$

where  $j'$  is a  $1 \times b$  vector. The elements of the vectors  $r$  and  $c$

are sometimes referred to as row and column masses.

We now convert each row and column of  $P$  to a

profile. The  $i$ th row profile  $r'_i$ ,  $i = 1, 2, \dots, a$ , is defined

by dividing the  $i$ th row of either Table 1 or 2 by its

marginal total:

$$r'_i = \left( \frac{p_{i1}}{p_i}, \frac{p_{i2}}{p_i}, \dots, \frac{p_{ib}}{p_i} \right) = \left( \frac{n_{i1}}{n_i}, \frac{n_{i2}}{n_i}, \dots, \frac{n_{ib}}{n_i} \right)$$

The elements in each  $r'_i$  are relative frequencies, and

therefore they sum to 1:

$$r'_i \cdot j = \sum_{j=1}^b \frac{n_{ij}}{n_i} = \frac{n_{i.}}{n_i} = 1$$

We define  $D_r$  as

$$D_r = \text{diag}(r) = \begin{bmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_a \end{bmatrix}$$

Similarly, the  $j$ th column profile  $c_j$ ,  $j = 1, 2, \dots, b$ , is

defined by dividing the  $j$ th column of either Table 1 or

Table 2 by its marginal total:

$$c_j = \left( \frac{p_{1j}}{p_{.j}}, \frac{p_{2j}}{p_{.j}}, \dots, \frac{p_{aj}}{p_{.j}} \right) = \left( \frac{n_{1j}}{n_{.j}}, \frac{n_{2j}}{n_{.j}}, \dots, \frac{n_{aj}}{n_{.j}} \right)'$$

elements in each  $c_j$  are relative frequencies, and

therefore they sum to 1:

$$j'c_j = \sum_{i=1}^a \frac{n_{ij}}{n_{.j}} = \frac{n_{.j}}{n_{.j}} = 1$$

By defining

$$D_c = \text{diag}(c) = \begin{bmatrix} p_{.1} & 0 & \dots & 0 \\ 0 & p_{.2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_{.b} \end{bmatrix}$$

Then the two matrices  $D_r$  and  $D_c$  can be transformed to matrix  $B$  by multiplying the matrix  $P$  as follows:

$$B = D_r^{-\frac{1}{2}} \times P \times D_c^{-\frac{1}{2}} \tag{2}$$

Since we are dealing with diagonal matrices, the operations  $D_r^{-\frac{1}{2}}$  and  $D_c^{-\frac{1}{2}}$  are a diagonal element with

$D_r^{-\frac{1}{2}}$  and  $D_c^{-\frac{1}{2}}$ , the off-diagonal zero elements in each matrix. The matrix  $B$  will be  $(a \times b)$ , with transformed element  $b_{ij}$  corresponding to every original element

$n_{ij}$ . The matrix of cross-products of the columns is simply:

$$C = B' \times B \tag{3}$$

Similarly, the matrix of cross-product of the rows is

$$R = B \times B' \tag{4}$$

The Eigen values of  $C$  and  $R$  will be identical, except that  $R$  will have  $(a-b)$  additional Eigenvalues, each of which will be zero. The Eigenvectors of  $C$  can be converted to correspondence factor loadings by multiplying each vector by its corresponding singular value which is the square root of the corresponding Eigenvalues.

That is:

$$L(C) = U \times A^{\frac{1}{2}} \tag{5}$$

where  $L(C)$  represents the factor loadings of column matrix. In the matrix notation we have used earlier, the singular values of  $C$  can be thought of as occurring along the diagonal of an  $(a \times a)$  square root of Eigenvalues matrix  $\sqrt{\lambda}$ , whose off-diagonal elements are all zero. Eigenvectors of  $C$  are from the column of an  $(a \times a)$  matrix,  $U$ . The matrix Equation is used to determine the loadings  $C$ .

The scores of each of the observations on the  $b$  correspondence factors are simply:

$$S(C) = B \times L(C) \tag{6}$$

We can calculate  $R$  correspondence loadings and scores, and then loadings are found by multiplying the elements of the Eigenvectors by the square root of the associated Eigenvalues

$$L(R) = V \times A^{\frac{1}{2}} \tag{7}$$

where  $L(R)$  represents the factor loadings of row matrix and  $V$  is the  $(b \times b)$  matrix whose columns contain the  $b$  Eigenvectors of  $R$  scores which are:

$$S(R) = B' \times L(R) \tag{8}$$

**Table 1** Contingency Table with a Rows and b Columns

		Columns				Row Total
		1	2	...	B	
Rows	1	$n_{11}$	$n_{12}$	...	$n_{1b}$	$n_{1.}$
	2	$n_{21}$	$n_{22}$	...	$n_{2b}$	$n_{2.}$
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	a	$n_{a1}$	$n_{a2}$	...	$n_{ab}$	$n_{a.}$
Column Total		$n_{.1}$	$n_{.2}$	...	$n_{.b}$	$N$

**Table 2** Correspondence Matrix of Relative Frequencies

		Columns				Row Total
		1	2	...	b	
Rows	1	$P_{11}$	$P_{12}$	...	$P_{1b}$	$P_{1.}$
	2	$P_{21}$	$P_{22}$	...	$P_{2b}$	$P_{2.}$
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	a	$P_{a1}$	$P_{a2}$	...	$P_{ab}$	$P_{a.}$
Column Total		$P_{.1}$	$P_{.2}$	...	$P_{.b}$	1

**5.0 SHORTCUT TRANSFORMATION BETWEEN C AND R MATRICES**

We use the number of calculations to find the relationship between factor loadings and factor scores for each of the column matrix C and row matrix R and try to find the matrix C of the matrix R and vice versa and as follows.<sup>1</sup>

$$L(R) = B \times L(C) \times A^{-1} \tag{9}$$

$$L(R) = S(C) \times A^{-1} \tag{10}$$

in which we found the matrix of factor loadings  $L(R)$  by using the matrix of factor scores  $S(C)$ , and in the same way we found the matrix of factor loadings  $L(C)$  by using the matrix of factor scores  $S(R)$ .

$$L(C) = B' \times L(R) \times A^{-1} \tag{11}$$

$$L(C) = S(R) \times A^{-1} \tag{12}$$

and rearrangement of the two formulas (10) and (12) and multiplying two sides with the square root of the Eigenvalues  $\frac{1}{A^2}$  we find that the factor scores matrix for each of the R and C as follows:

$$S(C) = L(R) \times A^{\frac{1}{2}} \tag{13}$$

$$S(R) = L(C) \times A^{\frac{1}{2}}$$

(14)

Thus, we found the R matrix using the C matrix and vice versa without using the constraint mentioned earlier.

**6.0 COLLECTION DATA**

Data used in this study were taken from the Ibn Sina Hospital for (30) female patients infected by Osteoporosis. Some of the variables could be influential on the disease and these variables are explained in Appendix A.

**7.0 STATISTICAL ANALYSIS**

Minitab Statistical Program has been used to find row & column matrices in the multiple correspondence analysis and the results are as as in the Tables (3), (4), (5), (6), (7) and (8) respectively:

The results corresponding to the column matrix (10×10) are reported in Table 3. We used Equation 3 to get on C matrix. R matrix is a very huge matrix (30×30) and we can get on it by using Equation 4.

The Eigenvalues for columns and rows matrices are reported in Table 4 and 5 respectively. We noticed that 7 Eigenvalues from 10 are selected to explain the total variance and 7 Eigenvalues from 30 in row matrix are selected to explain the total variance and the remaining values are zero. It is obvious from Eigenvalues results of column and row matrices that the results are equivalent and the number of selected Eigenvalues is 7.

The Eigenvectors for column matrix are presented in Table 6. Eigenvectors for row matrix are very huge

matrix (30×30) and we can get on it by using R matrix in Equation 4.

The factor scores for rows and column matrix are presented in Tables 7 and 10 in Appendix C. The factor loadings for column matrix are presented in

Tables 8. The factor loadings for row matrix are a very huge matrix (30×30) and we can get on it by using Equation 4. Finally the matrix B (30×10) is presented in Appendix B.

**Table 3** Column Matrix

0.187	0.098	0.124	0.121	0.132	0.128	0.157	0.095	0.152	0.122
0.098	0.056	0.067	0.067	0.072	0.068	0.086	0.053	0.081	0.068
0.124	0.067	0.092	0.091	0.091	0.088	0.108	0.069	0.104	0.086
0.121	0.067	0.091	0.148	0.087	0.090	0.101	0.073	0.103	0.088
0.132	0.072	0.091	0.087	0.098	0.094	0.117	0.071	0.109	0.091
0.128	0.068	0.088	0.090	0.094	0.094	0.112	0.070	0.105	0.088
0.157	0.086	0.108	0.101	0.117	0.112	0.142	0.085	0.129	0.109
0.095	0.053	0.069	0.073	0.071	0.070	0.085	0.056	0.078	0.068
0.152	0.081	0.104	0.103	0.109	0.105	0.129	0.078	0.130	0.102
0.122	0.068	0.086	0.088	0.091	0.088	0.109	0.068	0.102	0.092

**Table 4** Eigenvalues of Column Matrix

1	0.06	0.012	0.006	0.005	0.003	0.003	0.002	0.000	0.000
---	------	-------	-------	-------	-------	-------	-------	-------	-------

**Table 5** Eigenvalues of Row Matrix

1	0.06	0.012	0.006	0.005	0.003	0.003	0.002	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

**Table 6** Eigenvectors of Column Matrix

-0.424	-0.189	-0.484	-0.337	0.060	0.445	0.470	-0.072	0.085	0.036
-0.230	-0.070	0.047	0.050	-0.297	-0.364	0.380	0.657	-0.340	0.161
-0.296	-0.001	-0.001	0.774	0.364	0.235	0.033	-0.075	-0.255	0.232
-0.309	0.927	-0.055	-0.109	0.001	-0.110	0.056	-0.110	-0.042	-0.006
-0.309	-0.164	0.067	-0.049	0.125	-0.160	-0.061	-0.155	-0.433	-0.784
-0.301	-0.052	0.157	-0.418	0.329	0.186	-0.583	0.327	-0.202	0.276
-0.368	-0.239	0.310	-0.142	0.136	-0.529	0.142	-0.472	0.176	0.346
-0.230	0.040	0.433	0.111	0.215	0.129	0.152	0.377	0.644	-0.321
-0.352	-0.098	-0.534	0.244	-0.259	-0.286	-0.474	0.091	0.367	-0.068
-0.293	-0.036	0.403	0.061	-0.721	0.404	-0.135	-0.202	-0.056	0.037

**Table 7** Factor Scores of Row Matrix

-0.4241	0.0115	-0.0059	-0.0022	0.0003	0.0016	-0.0015	-0.0002	-0.0001	0.0000
-0.2298	0.0043	0.0006	0.0003	-0.0017	-0.0013	-0.0012	0.0014	0.0003	0.0001
-0.2957	0.0001	-0.0000	0.0051	0.0021	0.0008	-0.0001	-0.0002	0.0002	0.0001
-0.3094	-0.0563	-0.0007	-0.0007	0.0000	-0.0004	-0.0002	-0.0002	0.0000	0.0000
-0.3094	0.0099	0.0008	-0.0003	0.0007	-0.0006	0.0002	-0.0003	0.0004	-0.0004
-0.3013	0.0032	0.0019	-0.0027	0.0019	0.0007	0.0018	0.0007	0.0002	0.0001
-0.3678	0.0145	0.0038	-0.0009	0.0008	-0.0019	-0.0004	-0.0010	-0.0001	0.0002
-0.2298	-0.0024	0.0053	0.0007	0.0012	0.0005	-0.0005	0.0008	-0.0005	-0.0002
-0.3518	0.0060	-0.0065	0.0016	-0.0015	-0.0010	0.0015	0.0002	-0.0003	0.0000
-0.2929	0.0022	0.0049	0.0004	-0.0042	0.0014	0.0004	-0.0004	0.0000	0.0000

Table 8 Factor Loadings of Column Matrix

-0.424	-0.047	-0.053	-0.027	0.005	0.026	0.026	-0.003	0.002	0.001
-0.230	-0.017	0.005	0.004	-0.023	-0.022	0.021	0.030	-0.010	0.004
-0.296	-0.000	-0.000	0.063	0.028	0.014	0.002	-0.003	-0.007	0.005
-0.309	0.229	-0.006	-0.009	0.000	-0.007	0.003	-0.005	-0.001	0.000
-0.309	-0.040	0.007	-0.004	0.010	-0.009	-0.003	-0.007	-0.012	-0.017
-0.301	-0.013	0.017	-0.034	0.025	0.011	-0.033	0.015	-0.006	0.006
-0.368	-0.059	0.034	-0.011	0.010	-0.031	0.008	-0.022	0.005	0.008
-0.230	0.010	0.048	0.009	0.016	0.008	0.009	0.017	0.019	-0.007
-0.352	-0.024	-0.059	0.020	-0.020	-0.017	-0.027	0.004	0.011	-0.002
-0.293	-0.009	0.044	0.005	-0.055	0.024	-0.008	-0.009	-0.002	0.001

## 8.0 CONCLUSIONS

Multiple correspondence analysis is very important subject and it is used as a data analysis technique with ordered categorical and dichotomous data. Multiple correspondence analysis is used to find row and column matrix. In this paper, it is possible to shortcut transformation between the results of row and column matrix at any stage of analysis, and important results in the multiple correspondence analysis that can be found by going between the matrices are: Eigenvalues and Eigenvectors and factor loadings and factor scores are obtained from the results of the matrix  $R$  using Equations (9, 10, 14) respectively, and the results of the matrix  $C$  can be obtained from the Equations (11, 12, 13), respectively. The results of factor loadings for column matrix is that the variables  $X_1, X_4, X_5, X_6, X_7, X_9$ , are significant with opposite effect on the study, but the remaining of variables are not significant.

## Acknowledgement

We acknowledge the financial support from University Teknologi Malaysia for the Research University Grant (Q.J130000 .2526. 06H68) and the Ministry of Higher Education (MOHE) of Malaysia.

## References

- [1] J. C., D. 1986. *Statistics And Data Analysis In Geology*. John Wiley & Sons, New York.
- [2] Thanoon, T. Y., Adnan, R., & Saffari, S. E. 2014. Multiple Factor Analysis With Continuous And Dichotomous Variables. In *International Conference On Quantitative Sciences And Its Applications (ICOQSIA 2014): Proceedings of the 3rd International Conference on Quantitative Sciences and Its Applications*. AIP Publishing. 1635: 926-933.
- [3] Benzécri, J.-P. and Benzécri, F. 1980. *Analyse des Correspondances: exposé élémentaire*: Dunod.
- [4] Hill, M. O. 1974. Correspondence Analysis: A Neglected Multivariate Method. *Applied Statistics*. 340-354.
- [5] Lebart, L., Morineau, A. and Warwick, K. M. 1984. *Multivariate Descriptive Statistical Analysis: Correspondence Analysis And Related Techniques For Large Matrices*. New York: Wiley.
- [6] Greenacre, M. J. 1984. *Theory And Applications Of Correspondence Analysis*.
- [7] Khattree, R. and Naik, D. N. 1999. *Applied Multivariate Statistics With SAS Software*. SAS Institute.
- [8] Gower, J. C. and Hand, D. J. 1996. *Biplots*. London: Chapman & Hall.
- [9] Benzécri, J.-P. 1992. *Correspondence Analysis Handbook*. Marcel Dekker New York.
- [10] Heiser, W. J. 1987. Correspondence Analysis With Least Absolute Residuals. *Computational Statistics & Data Analysis*. 5(4): 337-356.
- [11] Hlávka, Z. 2007. *Multivariate Statistics: Exercises and Solutions*. Springer.
- [12] Hrdle, W. and Simar, L. 2007. *Applied Multivariate Statistical Analysis*. Springer.
- [13] Rencher, A. C. 2002. *Methods Of Multivariate Analysis*. New York: J. Wiley.
- [14] Du, Y. 2003. Multiple Correspondence Analysis in Marketing Research. *Department of Mathematics and Computer Science, Duquesne University*.
- [15] Abdi, H. and Valentin, D. 2007. Multiple Correspondence Analysis. *Encyclopedia Of Measurement And Statistics*. 651-657.
- [16] Ambrogi, F., Biganzoli, E. and Boracchi, P. 2005. Multiple Correspondence Analysis In S-PLUS. *Computer Methods And Programs In Biomedicine*. 79(2): 161-167.
- [17] Loslever, P., Laassel, E. M. and Angue, J. 1994. Combined Statistical Study Of Joint Angles And Ground Reaction Forces Using Component And Multiple Correspondence Analysis. *Biomedical Engineering, IEEE Transactions on*. 41(12): 1160-116.
- [18] Greenacre, M. and Blasius, J. 2006. *Multiple Correspondence Analysis And Related Methods*. CRC Press.
- [19] Le Roux, B. and Rouanet, H. 2010. *Multiple Correspondence Analysis*. Sage. 163.
- [20] Panagiotakos, D. B. and Pitsavos, C. 2004. Interpretation Of Epidemiological Data Using Multiple Correspondence Analysis And Log-Linear Models. *Journal of Data Science*. 2(1): 75-86.
- [21] Reis, A., Sousa, A., Ferreira da Silva, E., Patinha, C. and Fonseca, E. C. 2004. Combining Multiple Correspondence Analysis With Factorial Kriging Analysis For Geochemical Mapping Of The Gold-Silver Deposit At Marrancos (Portugal). *Applied Geochemistry*. 19(4): 623-631.



## APPENDICES

## Appendix A

## Data Description

- X<sub>1</sub>: age , 1:"30-39" , 2: "40-49" , 3 : "50-59" , 4: "60-69" , 5 : "70-79" , 6 : "80-89".  
 X<sub>2</sub>: Residence, 1: "urban", 2:"rural".  
 X<sub>3</sub>: Marital Status, which contains 1: single, 2: married.  
 X<sub>4</sub>: Level of education. 1:"illiterate" 2: "read and write" 3: "primary" 4: "middle" 5: "middle" 6: "Institute and more."  
 X<sub>5</sub>: Profession: 1: "Employee", 2: "housewife".  
 X<sub>6</sub>: bmi (body mass index), 1: "less or equal 30 ", 2: "more than 30 ".  
 X<sub>7</sub>: Number of Pregnancy. 1: "none", 2: "1-5", 3: "6 and more ".  
 X<sub>8</sub>: stopped the session, 1: "yes", 2: "no".  
 X<sub>9</sub>: any age stopped, 1"did not stop", 2: "49 years", 3: "50 years and more".  
 X<sub>10</sub>: previous fractures, 1: "yes", 2: "no".

## Appendix B

Table 9 B – Matrix

0.077	0.035	0.055	0.158	0.053	0.054	0.044	0.035	0.069	0.055
0.084	0.077	0.060	0.029	0.057	0.029	0.072	0.039	0.076	0.061
0.084	0.077	0.030	0.029	0.057	0.059	0.072	0.039	0.076	0.061
0.042	0.039	0.060	0.172	0.029	0.059	0.048	0.077	0.025	0.061
0.064	0.040	0.061	0.029	0.059	0.060	0.074	0.040	0.077	0.062
0.086	0.040	0.061	0.029	0.059	0.060	0.074	0.040	0.077	0.031
0.064	0.040	0.061	0.029	0.059	0.060	0.074	0.040	0.077	0.062
0.084	0.039	0.030	0.086	0.057	0.059	0.072	0.039	0.050	0.061
0.084	0.039	0.060	0.057	0.057	0.059	0.072	0.039	0.076	0.030
0.064	0.040	0.061	0.059	0.059	0.060	0.049	0.040	0.077	0.062
0.066	0.041	0.032	0.090	0.060	0.062	0.076	0.041	0.053	0.032
0.063	0.039	0.060	0.172	0.029	0.029	0.024	0.039	0.076	0.061
0.086	0.040	0.031	0.029	0.059	0.060	0.074	0.040	0.077	0.062
0.107	0.040	0.061	0.029	0.059	0.060	0.049	0.040	0.077	0.031
0.105	0.039	0.060	0.029	0.057	0.059	0.072	0.039	0.050	0.061
0.066	0.041	0.063	0.030	0.060	0.062	0.076	0.041	0.053	0.064
0.088	0.041	0.032	0.030	0.060	0.062	0.076	0.041	0.053	0.064
0.045	0.042	0.065	0.031	0.062	0.064	0.078	0.083	0.027	0.065
0.086	0.040	0.061	0.029	0.059	0.030	0.074	0.040	0.077	0.062
0.080	0.037	0.057	0.082	0.055	0.056	0.069	0.037	0.072	0.058
0.086	0.040	0.061	0.029	0.059	0.060	0.074	0.040	0.077	0.031
0.068	0.042	0.032	0.031	0.062	0.064	0.078	0.042	0.054	0.065
0.082	0.038	0.059	0.056	0.056	0.057	0.071	0.038	0.074	0.059
0.086	0.040	0.061	0.029	0.059	0.060	0.074	0.040	0.052	0.062
0.090	0.042	0.065	0.031	0.062	0.032	0.078	0.042	0.054	0.033
0.088	0.041	0.032	0.030	0.060	0.062	0.076	0.041	0.053	0.064
0.082	0.038	0.059	0.084	0.056	0.057	0.071	0.038	0.074	0.030
0.105	0.039	0.060	0.029	0.057	0.059	0.048	0.039	0.076	0.061
0.066	0.041	0.063	0.090	0.060	0.062	0.051	0.041	0.053	0.032
0.046	0.043	0.067	0.032	0.064	0.033	0.080	0.043	0.056	0.067

## Appendix C

Table 10 Factor Scores of Column Matrix

-0.2031	0.0246	-0.0020	-0.0005	-0.0004	0.0001	-0.0002	-0.0003	-0.0002	-0.0001
-0.1862	-0.0076	-0.0008	0.0016	-0.0019	-0.0008	0.0013	0.0006	-0.0001	0.0000
-0.1862	-0.0080	-0.0002	-0.0012	-0.0019	-0.0009	0.0003	0.0011	-0.0001	0.0001
-0.1862	0.0315	0.0047	0.0001	0.0007	0.0005	0.0005	0.0005	0.0002	0.0002
-0.1817	-0.0065	0.0007	0.0011	-0.0003	-0.0002	-0.0010	-0.0001	0.0000	0.0001
-0.1817	-0.0073	-0.0018	0.0003	0.0015	-0.0004	-0.0002	0.0001	0.0001	0.0001
-0.1817	-0.0065	0.0007	0.0011	-0.0003	-0.0002	-0.0010	-0.0001	0.0000	0.0001
-0.1862	0.0064	0.0007	-0.0024	-0.0006	0.0000	0.0003	-0.0004	0.0000	0.0000
-0.1862	-0.0005	-0.0020	0.0001	0.0014	-0.0006	-0.0001	0.0000	0.0001	0.0001
-0.1817	0.0016	-0.0003	0.0011	-0.0006	0.0003	-0.0011	0.0003	-0.0001	-0.0001
-0.1771	0.0080	0.0005	-0.0021	0.0010	-0.0014	0.0000	-0.0001	-0.0000	-0.0001
-0.1862	0.0307	-0.0026	0.0014	-0.0018	0.0003	0.0001	0.0000	0.0001	0.0000
-0.1817	-0.0075	-0.0005	-0.0014	-0.0011	-0.0001	-0.0005	0.0000	0.0003	-0.0001
-0.1817	-0.0068	-0.0038	0.0000	0.0013	0.0009	0.0001	0.0006	0.0001	-0.0001
-0.1862	-0.0077	-0.0001	-0.0006	0.0004	0.0013	0.0008	-0.0003	-0.0001	0.0001
-0.1771	-0.0061	0.0023	0.0006	0.0002	0.0002	-0.0004	-0.0002	-0.0003	0.0001
-0.1771	-0.0071	0.0011	-0.0020	-0.0006	0.0004	0.0002	-0.0001	0.0000	0.0000
-0.1723	-0.0041	0.0071	0.0010	0.0013	0.0004	0.0001	0.0005	0.0002	-0.0002
-0.1817	-0.0071	-0.0010	0.0015	-0.0010	0.0000	0.0005	-0.0006	0.0003	-0.0001
-0.1948	0.0055	-0.0008	0.0000	-0.0002	0.0000	-0.0003	-0.0004	0.0000	0.0001
-0.1817	-0.0073	-0.0018	0.0003	0.0015	-0.0004	-0.0002	0.0001	0.0001	0.0001
-0.1723	-0.0062	0.0023	-0.0014	-0.0007	-0.0002	-0.0004	-0.0001	0.0000	0.0000
-0.1905	-0.0008	-0.0006	0.0002	-0.0002	0.0001	-0.0004	-0.0003	0.0000	0.0001
-0.1817	-0.0069	0.0011	0.0000	0.0003	0.0008	0.0002	-0.0003	-0.0002	0.0001
-0.1723	-0.0066	-0.0009	0.0009	0.0013	-0.0003	0.0015	-0.0004	0.0000	-0.0001
-0.1771	-0.0071	0.0011	-0.0020	-0.0006	0.0004	0.0002	-0.0001	0.0000	0.0000
-0.1905	0.0059	-0.0021	-0.0002	0.0014	-0.0008	0.0000	-0.0001	0.0001	0.0001
-0.1862	-0.0069	-0.0024	0.0002	-0.0004	0.0016	-0.0001	0.0003	0.0000	-0.0001
-0.1771	0.0095	-0.0004	0.0002	0.0017	-0.0002	-0.0001	0.0004	-0.0004	-0.0001
-0.1675	-0.0049	0.0031	0.0023	-0.0007	-0.0007	0.0001	-0.0006	-0.0001	-0.0001