



# Asian Journal of Mathematics & Statistics

ISSN 1994-5418

### Single Ordinal Correspondence Analysis with External Information\*

<sup>1</sup>Amenta Pietro, <sup>1</sup>Simonetti Biagio and <sup>2</sup>Beh Eric

<sup>1</sup>Università del Sannio

<sup>2</sup>University of Western Sydney, Australia

---

**Abstract:** Several non-iterative procedures for performing correspondence analysis with external information have been proposed in literature. The interpretation of the multidimensional representation of the row and column categories may be greatly simplified if additional information about the row and column structure are incorporated. In this paper, a new combined approach to impose external information (as linear constraints) in analyzing a contingency table which can be of an ordinal nature, is showed. Linear constraints are imposed using the polynomial approach to correspondence analysis. The classical approach to correspondence analysis decomposes the Pearson chi-squared statistic into singular values by partitioning the matrix of Pearson contingencies using singular value decomposition. The polynomial approach to correspondence analysis decomposes the same statistic by partitioning the matrix of Pearson contingencies using orthogonal polynomials rather than singular value decomposition. An alternative approach to partitioning the Pearson chi-squared statistic for a two-way contingency table is essentially to combine the approach of orthogonal polynomials for the ordered columns and singular vectors for the unordered rows. With this mixed approach, the researcher can determine any statistically significant sources of variation (location, dispersion and higher order components) of the ordered columns along the a particular axis using the simple correspondence analysis. Main aim of the present study is to introduce external information to this approach. In our proposal external information, such as taking into account that categories are not equally spaced is then included directly on suitable matrices which reflect the most important components. This approach allows for one to overcome the problem of having to impose linear constraints at the variables based on subjective decisions.

**Key words:** Ordinal correspondence analysis, external information, orthogonal polynomials

### INTRODUCTION

Canonical correspondence analysis is a widely used tool for obtaining a graphical representation of the dependence structure between the rows and columns of a contingency table (Benzecri, 1980; Greenacre, 1984; Lebart *et al.*, 1984; Nishisato, 1980). This procedure can also be used as a graphical analytic tool in sensory analysis. For example McEwan and Schlich (1992) used correspondence analysis to describe the consumer preference attributes of eight commercially available strawberry jam. The graphical representation is achieved by assigning scores in the form of coordinates to the row and column categories yielding a correspondence plot. Generally simple correspondence analysis is performed by applying a singular value decomposition to the standardized residuals of a two-way contingency table. Many, including Greenacre (1984), discuss in detail the theoretical, computational and application issues of the analysis. This decomposition ensures that the maximum information regarding the association between the two categorical variables are accounted for in the first

---

**Corresponding Author:** Simonetti Biagio, Università del Sannio

*Originally Published in Asian Journal of Mathematics and Statistics, 2008*

one or two dimensions of a correspondence plot. However, while such a plot can identify those categories that are similar, for those that are different, it is not as easy to clarify how the categories are different.

The interpretation of the relationship between the categories of a two-way, or multi-way contingency table may be greatly simplified if additional information concerning the row and column structure of the table is available. By incorporating this external information through linear constraints on the row and/or column scores a representation of the data may be obtained that is not only more parsimonious but is also easier to understand. In the classical analysis, Böckenholt and Böckenholt (1990) considered this problem.

In order to overcome both of these remarks, the aim of this study is to consider an extension of Simple correspondence analysis of ordinal cross-classifications using orthogonal polynomials (Beh, 1997) which takes into account external information through linear/ordinal constraints on the row and/or column scores.

In Section 3 we analyze data that is derived from a sensory survey. The survey was conducted on a panel of 300 consumers who were asked to evaluate the organoleptic characteristics of a typical cheese of south of Italy, "Mozzarella". The results obtained shown that by imposing constraints on ordinal correspondence analysis it is possible to improve the interpretation of the relationship between the categories.

### A RESTRICTED APPROACH FOR SINGLE ORDINAL CORRESPONDENCE ANALYSIS

Correspondence analysis is a popular graphical tool used to analyze contingency tables. It is a technique that has, in the last decade, experienced a boom in the diversity of applications and adaptations and has been exposed to most disciplines. In the past, it has commonly been performed by applying a singular value decomposition to a transformation of the data in the contingency table.

Let  $P$  be an  $I \times J$  two-way contingency table describing the joint distribution of two categorical variables where the  $(i,j)$ th cell entry is given by the proportion  $p_{ij} = n_{ij}/n$  for  $(i = 1, \dots, I)$  and  $(j = 1, \dots, J)$  with  $\sum_{i=1}^I \sum_{j=1}^J p_{ij} = 1$  and  $n = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$ . Let  $D_i$  be the diagonal matrix whose elements are the row masses  $p_{i+} = \sum_{j=1}^J p_{ij}$  and let  $D_j$  be the diagonal matrix whose elements are the column masses  $p_{+j} = \sum_{i=1}^I p_{ij}$ . Correspondence analysis is performed by applying singular value decomposition to the matrix of standardized residuals  $\Pi = D_i^{-1/2}(P - D_i 11^T D_j)D_j^{-1/2}$ .

According to the principle of Restricted Eigenvalue Problem (Rao, 1973), Böckenholt and Böckenholt (1990) proposed a canonical analysis of contingency tables (RCCA) which takes into account additional information about the row and column categories of the table.

Let  $H$  and  $G$  be the matrices of external information (as linear restrictions) of order  $(I \times K)$  and  $(J \times L)$  and ranks  $K$  and  $L$ , respectively, such that  $H^T X = 0$  and  $G^T Y = 0$  where  $X$  and  $Y$  are the restricted standardized row and column scores. Restricted CCA scores (Böckenholt and Böckenholt, 1990) are obtained by a SVD of the matrix

$$\{(I - D_i^{-1/2} H (H^T D_i^{-1} H)^{-1} H^T D_i^{-1/2}) \Pi (I - D_j^{-1/2} G (G^T D_j^{-1} G)^{-1} G^T D_j^{-1/2})\} = U \Theta V^T$$

where,  $U^T U = I$ ,  $V^T V = I$  and  $\Theta$  is a diagonal matrix with eigenvalues  $\eta$  in descending order. Standardized row and column scores are given by  $X = D_i^{-1/2} U$  and  $Y = D_j^{-1/2} V$ , respectively, such that  $X^T D_i X = I$ ,  $Y^T D_j Y = I$ ,  $I^T D_i X = 0$  and  $I^T D_j Y = 0$ . Constraints are often imposed by making use of orthogonal polynomials which are suitable for subdividing total variation of the scores into linear, quadratic, cubic, etc., components. For example, in order to obtain a linear order for the standard scores, authors eliminate the effects of the quadratic and cubic trend by including suitable constraint matrices.

The classical approach to correspondence analysis is obtained when  $h = [D_1 | 1]$  and  $G[D_j | 1]$  (absence of external information). For deeper remarks on linear constraints in correspondence analysis see also Böckenholt and Böckenholt (1990), Takane *et al.* (1991), Böckenholt and Takane (1994) and Hwang and Takane (2002).

It is possible to consider external information directly on suitable matrices which reflect the different most important components by taking into account the Pearson chi-squared partition obtained by using the orthogonal polynomial for contingency tables (Emerson, 1968). The advantage of using these orthogonal polynomials relies in the fact that the order information is considered in the analysis and the resulting scoring scheme permits a clear interpretation of the linear, quadratic or higher order trend components.

The correspondence analysis approach of Beh (1997) decomposes the  $(i, j)$ 'th Pearson ratio so that

$$\alpha_{ij} = 1 + \sum_{u=1}^{I-1} \sum_{v=1}^{J-1} a_u(i) b_v(j) \lambda_{uv}$$

For this method of decomposition,  $\lambda_{uv}$  is the  $(u, v)$ 'th generalised correlation (Davy *et al.*, 2003) while  $\{a_u(i) : u = 1, \dots, I-1\}$  and  $\{b_v(j) : v = 1, \dots, J-1\}$  are the orthogonal polynomials for the  $i$ 'th row and  $j$ 'th column respectively. The polynomials are calculated using the general recurrence relation of Emerson (1968) and require a set of scores to reflect the nature of the variables. For example, the polynomials for ordinal variables can be constructed using natural scores.

This approach to correspondence analysis uses the Bivariate Moment Decomposition (BMD) to identify linear (location), quadratic (dispersion) and higher order moments for each of the ordinal variables. This feature is not readily available by using classical singular value decomposition. The first axis of the correspondence plot reflects differences in location and the second axis reflects differences in dispersion. Higher order components are reflected by higher dimensional axes.

As remarked before, the general recurrence relation of Emerson's (1968) can be used only when a set of scores for the ordered column category has been considered. For sake of simplicity, only natural column scores,  $\{s_j(j) = j; j = 1, \dots, J\}$ , will be considered in this paper. Different scoring schemes can have distinct effects on the orthogonal polynomials and simple correspondence analysis (Beh, 1998).

Suppose that  $P$  has nominal row and ordered column categories. The set of column orthogonal polynomials are placed into a  $(J \times J)$  matrix,  $B$ , whose  $(v, j)$ 'th element is denoted by  $b_{vj}$  such that  $B^T D_j B = 1$ . Based on the recurrence relation  $b_{j0}$  is the first column of  $B$  whose elements are all equal to 1 (trivial orthogonal polynomial) and it is assumed that when  $v = -1$ ,  $b_{jv} = 0$ .

There are various papers that deal with the partition of the symmetric chi-squared statistic with one or two ordered categories (Best and Rayner, 1996). An alternative approach to partitioning the Pearson chi-squared statistic for a two-way contingency table, which is described in Beh (2001), is to essentially combine the approach of orthogonal polynomials for the ordered columns and singular vectors for the unordered rows, so that

$$X^2 = \sum_{u=1}^{M^*} \sum_{v=1}^{J-1} Z_{(u)v}^2 \tag{1}$$

with  $M^* \leq I-1$  and where

$$Z_{(u)v} = \sqrt{n} \sum_{i=1}^I \sum_{j=1}^J p_{ij} a_{iu} b_{vj} \tag{2}$$

which are asymptotically standard normally distributed random variables. The parentheses around  $u$  indicates that (1) - (2) is concerned with a non-ordered set of row categories. Equation 2 can alternatively written in matrix form as

$$Z = \sqrt{n}A^T P B \tag{3}$$

where, A is the  $I \times (I-1)$  matrix of left singular vectors, while B is the  $J \times (J-1)$  matrix of the J-1 non-trivial column orthogonal polynomials. Previous equation can be re-arranged so that

$$D_i^{-1} P D_j^{-1} = A \left( \frac{Z}{\sqrt{n}} \right) B^T$$

The value of  $Z_{(u)v}$  means that each principal axis from a simple correspondence analysis can be partitioned into column component values. In this way, the researcher can determine the dominant source of variation of the ordered columns along a particular axis using the simple correspondence analysis.

Therefore the Pearson ratio of the singly ordered analysis is

$$\alpha_{ij} = \sum_{u=0}^{M^*} \sum_{v=0}^{J-1} a_{iu} \left( \frac{Z_{(u)v}}{\sqrt{n}} \right) b_v(j) \tag{4}$$

where,  $Z_{(u)v}$  is defined by (2). The value of  $a_{iu}$  is akin to the singular vector of simple correspondence analysis calculated using a singular value decomposition and so is not an orthogonal polynomial. It is associated with the row profiles, while the set of orthogonal polynomials,  $\{b_v(j)\}$ , are associated with the ordered columns.

Eliminating the trivial solution to (4) yields the Pearson contingencies

$$\alpha_{ij} - 1 = \sum_{u=1}^{M^*} \sum_{v=1}^{J-1} a_{iu} \left( \frac{Z_{(u)v}}{\sqrt{n}} \right) b_v(j) \tag{5}$$

leading to

$$\Pi = D_i^{-1/2} (P - D_i 1 1^T D_j) D_j^{-1/2} = A \left( \frac{Z}{\sqrt{n}} \right) B_*^T$$

where  $B_*$  is the matrix B without the first column  $b_0$ .

There are several properties that Beh (2001) derived which show the relationship between the singular values and the bivariate moments. Main properties are:

The row component associated with the  $m$ 'th principal axis is just the value of the  $m$ 'th largest eigenvalue. In the same way, total inertia may be written in terms of bivariate moments or as eigenvalues, such that

$$\frac{X^2}{n} = \sum_{u=1}^{M^*} \sum_{v=1}^{J-1} \frac{Z_{(u)v}^2}{n} = \sum_{m=1}^{M^*} \lambda_m^2 \tag{6}$$

so that

$$\lambda_m^2 = \sum_{v=1}^{J-1} \frac{Z_{(m)v}^2}{n} \tag{7}$$

- This method allows for a decomposition of the singular values into location, dispersion and higher order components. In this way, it can be applied to a non-ordered contingency table. That is, each singular value can be partitioned so that information concerning the mean difference and the spread of profiles can be found. Higher order moments can also be determined from such a partition.
- The row component values are arranged in descending order.
- For such a singly ordered analysis, it allows for the principal inertia of a simple correspondence plot to be partitioned into bivariate moments. When the principal inertia of the m'th principal axis is the sum of squares of the bivariate moments. For example, the largest eigenvalue may be calculated by considering the linear-by-linear, linear-by-quadratic and higher order bivariate moments, such that

$$\lambda_1^2 = \frac{Z_{(1)1}^2}{n} + \frac{Z_{(1)2}^2}{n} + \dots + \frac{Z_{(1)J-1}^2}{n} \tag{8}$$

- It is possible to identify which bivariate moment contributes the most to a particular eigenvalue and hence principal axis.

For other properties and an extension of this approach to double ordered contingency tables see Beh (1998, 1999, 2001).

It is therefore possible to use these results extend the Böckenholt and Böckenholt approach for a two-way cross-classification with nominal and ordinal categorical variables.

Let  $\Pi_v^* = A(\frac{Z}{n})B_v^T$  be the matrix identifying the most important linear (location), quadratic (dispersion) or higher order moments by using the previous results.

Restricted Single Ordinal Correspondence Analysis (RSOCA) scores are then obtained by performing a SVD of the matrix

$$\{I - D_1^{-1/2}H(H^T D_1^{-1}H)^{-1}H^T D_1^{-1/2}\} \Pi_v^* \{I - D_j^{-1/2}G(G^T D_j^{-1}G)^{-1}G^T D_j^{-1/2}\} = U\Theta V^T$$

where  $U^T U = I$ ,  $V^T V = I$  and  $\Theta$  is a diagonal matrix with eigenvalues  $\eta$  that are arranged in descending order.

Standardized row and column scores are given by  $X = D_1^{-1/2}U$  and  $Y = D_j^{-1/2}V$ , respectively, such that  $X^T D_1 X = I$ ,  $Y^T D_j Y = I$ ,  $I^T D_1 X = 0$  and  $I^T D_j Y = 0$ .

We remark that Single Ordinary Correspondence Analysis is obtained for  $H = [D_1 | 1]$  and  $G = [D_j | 1]$  (where there is an absence of external information).

By considering this approach it is possible to consider external information directly on suitable matrices that reflect the most important components. For example, researchers can impose external constraints in order to restrict the spacing to be the same for some modalities without considering those for ensuring linear order if they perform the restricted analysis over  $\Pi_v^*$  for  $v = 1$ . Moreover, they will be sure to eliminate the effects of the others trends working only on the component of main importance and interest.

This approach can be easily extended to the other versions of singly or doubly ordinal correspondence analysis (Beh, 1998, 1999, 2001)) as well as to the two and three way ordinal non-symmetrical approaches (D'Ambra *et al.*, 2002; Lombardo *et al.*, 2007; Beh *et al.*, 2007).

### APPLICATION

To demonstrate the applicability of the technique described above consider Table 1. Three dairy farms in Campania (Italy) participated in a study where 300 consumers were asked to evaluate their Overall Satisfaction (OS) of a given piece of mozzarella cheese. This was measured on a four point scale where the color (Milk, Cream, Ivory) of the cheese was considered. The 300 independent individuals that are classified into Table 1. While it is apparent that Table 1 may be considered as a 3×4×3 contingency table, in the spirit of the analysis of Böckenholt and Böckenholt (1990) we shall analyze it as a 9×4 table.

Table 2 summarizes the decomposition of the total inertia of Table 1 into location, dispersion and an error component. This error component consists of those measures of moment higher than the dispersion. It shows that the most dominant and only significant source of variation in the Satisfaction levels is in terms of their location, while the spread of the profiles is not a dominant source of variation.

To study the relationship between the two variables, a singly ordered correspondence analysis is performed. To do so the overall satisfaction categories are treated as ordinal and the row categories are nominal.

The representation in Fig. 1 shows the relationship between the modalities of the color (Fig. 1a) related to the three dairy-farms, Salernitani (D1), La Baronía (D2), Prati del Volturno (D3) and the modalities of the overall satisfaction (Fig. 1b). The first factorial plane is obtained by taking into account the linear component, ie first non-trivial orthogonal polynomial, from the decomposition (4). It explains approximately the 90% of the total inertia of the Table 1 (61% axis 1 and 29% axis 2). The factorial plane shows that customers have expressed a relatively low evaluation of the mozzarella product from the dairy-farm La Baronía with respect to all three cheese colors. The milk colored cheeses at the other two farms have a relatively high level of satisfaction. We also note that there is a strong association between the satisfaction modalities 3 and 4. Following this consideration we impose a linear constraint on the column modalities (overall satisfaction) with the aim to enforce a separation

Table 1: Cross-classification of the evaluation of mozzarella cheese

Dairy-farm	Color	Overall satisfaction 1(bad) - 4 (good)			
		1	2	3	4
Salernitani	Milk	8	6	8	9
	Cream	6	12	7	10
	Ivory	6	6	10	12
La Baronía	Milk	20	9	5	5
	Cream	10	6	6	8
	Ivory	9	7	6	9
Prati del Volturno	Milk	6	5	5	7
	Cream	5	11	8	7
	Ivory	5	12	13	16

Table 2: Decomposition of the total inertia  $X^2/n$

Components	Column inertia		
	Value	Contribution	p-value
Location	24.90485	80.13	0.00002
Dispersion	5.42294	17.47	0.14332
Error	0.74358	2.40	0.53724
Total inertia	31.07136	100.00	0.00192

between the low values of the overall satisfaction levels (1 and 2) and the high satisfaction levels (3 and 4). Thus a constrained solution is computed by setting  $G = D_11$  and  $H = (D_c1 | G_1)$  with  $G_1 = (0,0,-1,1)$ .

Under the restricted singly ordered correspondence analysis approach Fig. 2 gives the plot of the coordinates for the first two axes. The RSOCA plots allow one to better understand the relationship between the two variables. The plot on left-hand-side (Fig. 2a) shows that there is a strong association between the cream modalities of the three dairy-farms and by observing their proximity with the points in Fig. 2b, they are associated with the level OS 2 of overall satisfaction. Ivory colored cheese at Salernitani (D1) and Prati del Volturno (D3) appear to be similarly distributed by observing their proximity to one another. Their distance from any other point indicates that their location is different to many other dairy farm/color combinations This is true also for milk colored cheese at La Baronía.

The highest level of satisfaction is related to the Ivory modalities of the mozzarella products by the dairy-farms D1 and D3 according to the raw data of Table 1. The plot shows also that the linear constraint imposed allows the user to gain a better understand the differences between the high and low levels of satisfaction relating to the color of the cheese and the dairy-farms D1 and D3.

In order to better compare the ordinal and classical (Böckenholt and Böckenholt, 1990) approaches to the constrained analysis, restricted canonical correspondence analysis is performed with the following column constraint:

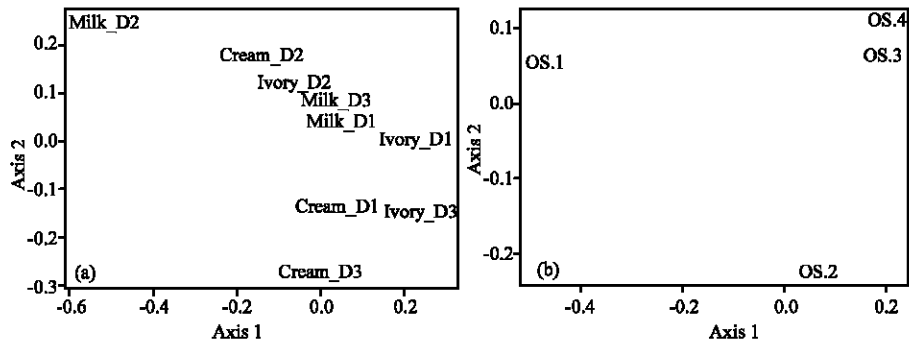


Fig. 1: Single ordinal correspondence analysis of mozzarella data

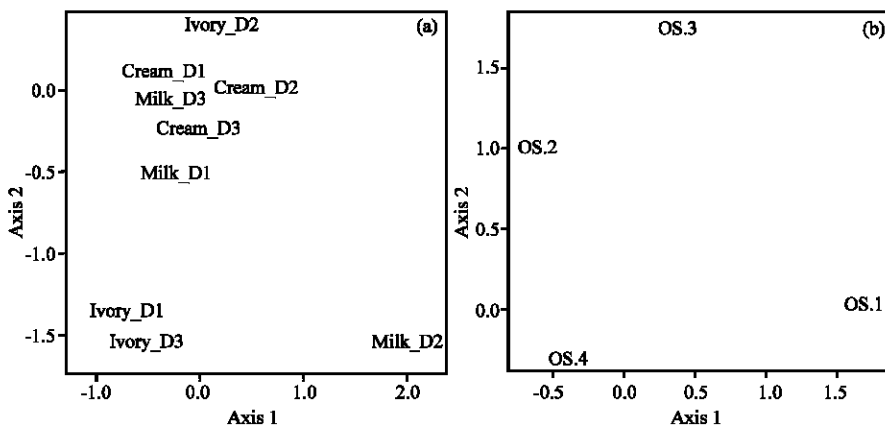


Fig. 2: Restricted single ordinal correspondence analysis of mozzarella data



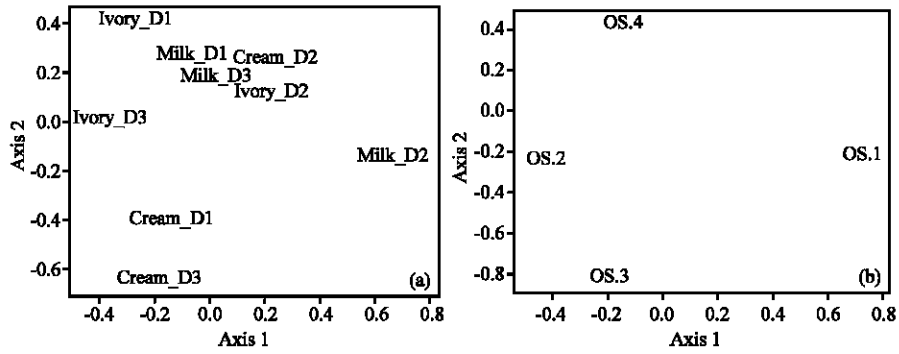


Fig. 3: Restricted (Böckenholt and Böckenholt) correspondence analysis of mozzarella data

$$G'_1 = \begin{vmatrix} -3 & -1 & 1 & 3 \\ 0 & 0 & -1 & 1 \end{vmatrix}$$

While the first row of the  $G'_1$  matrix ensures the linear spacing of the standardized columns (not necessary in RSOCA because only the linear component is taken into account), the second row aims to restrict the space between the low values of the overall satisfaction variable (1 and 2) and the high evaluation (3 and 4), as before defined for RSOCA.

By comparing Fig 2 with Fig 3, we can appreciate the different interpretation between the two sets of axes. While in RSOCA plot Ivory colored cheese at Salernitani (D1) and Prati del Volturno (D3) appear to be similarly distributed by observing their proximity to one another, they seem to be a little bit different from one another in Fig. 3, however with the classical restricted approach is not clear what the cause of this difference is. By using RSOCA we know that the difference is not in the first moment (location)

Figure 3 shows also that milk colored mozzarella cheese from La Baronía is very different compared to other colored cheeses from other dairy farms. RSOCA plot shows that this difference is due to its location being very different from other areas.

Finally, Fig. 3 highlights that “Cream” colored cheese at Salernitani (D1) and Prati del Volturno (D3) appear to be associated with the level OS 3 of overall satisfaction while by RSOCA we know that, according also to the raw data, they can be associated to OS 2.

### CONCLUSIONS

Several authors (Nishisato, 1980; Böckenholt and Böckenholt, 1990) highlighted that introducing linear constraints on the row and column coordinates of a correspondence analysis, representation may be greatly simplify the interpretation of the data matrix. The alternative approach of Beh for ordinal categorical data has been shown to be a more useful and informative correspondence analysis method than the classical technique commonly used. When Beh (1997) considered the partition of the chi-squared statistic for a two-way contingency table by combining the approach of orthogonal polynomials for the ordered columns and singular vectors for the unordered rows, he restricted the analysis to the integer valued scores. The problem with such a scoring scheme is that it assumes that the ordered categories are equally spaced. In general we know that this may not be the case.

Following our proposal, suitable linear constraints can be introduced in the combined Beh’s analysis in order to take into account not-equally spaced categories.

Moreover, in order to obtain a linear order for the standard scores, Böckenholt and Böckenholt remove the effects of the quadratic and cubic trend by including suitable constraint matrices even if

they do not know if they are statistically significant sources of variation. The knowledge of their significant allows researcher to improve the interpretation of the data matrix using suitable constraints for these significant components otherwise ignored. Finally, we consider linear constraints directly on suitable matrices that reflect only the most important components. Therefore, the proposed technique, could be usefully applied when, analyzing contingency table, the modality are non equi-spaced and there are the data are infected by non linear sources of variability.

## REFERENCES

- Beh, E.J., 1997. Simple correspondence analysis of ordinal cross-classifications using orthogonal polynomials. *Biometrical J.*, 39: 589-613.
- Beh, E.J., 1998. A comparative study of scores for correspondence analysis with ordered categories. *Bioemetrical J.*, 40: 413-429.
- Beh, E.J., 1999. Correspondence analysis of ranked data. Theory and methods, *Commun. Stat.*, 28: 1511-1533.
- Beh, E.J., 2001. Partitioning Pearson's chi-squared statistic for singly ordered two-way contingency tables. *Aust. New Zealand J. Stat.*, 43: 327-333.
- Beh, E.J., B. Simonetti and L. D'Ambra, 2007. Partitioning a non-symmetric measure of association for three-way contingency tables. *J. Multivariate Anal.*, 98: 1391-1411.
- Benzecri, J.P., 1980. *Pratique De L'analyse Des Donnees*. Vol. 1-3. Dunod, Paris.
- Best, D.J. and J.C.W. Rayner, 1996. Nonparametric analysis for doubly ordered two-way contingency tables. *Biometrics*, 52: 1153-1156.
- Böckenholt, U. and I. Böckenholt, 1990. Canonical analysis of contingency tables with linear constraints. *Psychometrika*, 55: 633-639.
- Böckenholt, U. and Y. Takane, 1994. Linear Constraints in Correspondence Analysis. In: *Correspondence Analysis in the Social Sciences: Recent Developments and Applications*. Greenacre, M.J. and J. Blasius (Eds.), Academic Press, New York, pp: 112-127.
- D'Ambra, L., R. Lombardo and P. Armenta, 2002. Non-symmetric Correspondence Analysis for Ordered Two-way Contingency Table. Invited Speaker XLI Riunione Scientifica della Società Italiana di Statistica, Milano.
- Davy, P.J., J.C.W. Rayner and E.J. Beh, 2003. Generalised Correlations and Simpson's Paradox. In: *Current Research in Modelling, Data Mining and Quantitative Techniques*. Pemajayantha, V., R.W. Mellor, S. Peiris and J.R. Rajasekera (Eds.), pp: 63-73.
- Emerson, P.L., 1968. Numerical construction of orthogonal polynomials from a general recurrence formula. *Biometrics*, 24: 696-701.
- Greenacre, M.J., 1984. *Theory and Applications of Correspondence Analysis*. Academic Press, New York.
- Hwang and Takane, 2002. Generalized structured component analysis. *Psychometrika*, 69: 81-99.
- Lebart, L., A. Morineau and K.M. Warwick, 1984. *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices*. John Wiley and Sons, New York.
- Lombardo, R., E.J. Beh and L. D'Ambra, 2007. Non-symmetric correspondence analysis for doubly ordinal contingency tables. *Comput. Stat. Data Anal.* (In Press).
- McEwan, J.A. and P. Schlich, 1992. Correspondence analysis in sensory evaluation. *Food Quality Preference*, 3: 23-36.
- Nishisato, S., 1980. *Analysis of Categorical Data: Dual Scaling and its Applications*. University of Toronto Press, Toronto.
- Rao, C.R., 1973. *Linear Statistical Inference and its Application*. John Wiley and Sons, New York.
- Takane, Y., H. Yanai and S. Mayakawa, 1991. Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika*, 56: 667-684.