Dimensionality reduction methods for contingency tables with ordinal variables

Metodi di riduzione delle dimensionalità per tabelle di contingenza con variabili ordinali

Luigi D'Ambra, Pietro Amenta, Antonello D'Ambra

Abstract Correspondence analysis is a widely used tool for obtaining a graphical representation of the interdependence between the rows and columns of a contingency table, by using a dimensionality reduction of the spaces. The maximum information regarding the association between the two categorical variables is then visualized allowing to understand its nature. Several extensions of this method take directly into account the possible ordinal structure of the variables by using different dimensionality reduction tools. Aim of this paper is to present an unified theoretical framework of several methods of correspondence analysis with ordinal variables.

Abstract *L'analisi delle corrispondenze ottiene una rappresentazione grafica della interdipendenza tra le righe e colonne di una tabella di contingenza mediante una riduzione della dimensionalità degli spazi. La massima informazione riguardante l'associazione tra le due variabili viene allora visualizzata in un sottospazio di dimensione ridotta per coglierne la natura. Diverse estensioni di tale approccio considerano anche la possibile struttura ordinale delle variabili, utilizzando diversi metodi di riduzione della dimensionalità. Scopo di questo lavoro è quello di presentare un quadro teorico unitario di alcuni estensioni dell'analisi delle corrispondenze in presenza di variabili ordinali.*

Key words: Contingency table, ordinal categorical variable, association index, singular value decomposition, bivariate and hybrid moment decomposition.

Luigi D'Ambra
Department of Economics, Management and Institutions, University of Naples, Italy, Via Mezzocannone 8, Napoli, e-mail: dambra@unina.it

Pietro Amenta
Department of Law, Economics, Management and Quantitative Methods, University of Sannio, Italy, Piazza Arechi II, Benevento, e-mail: amenta@unisannio.it

Antonello D'Ambra
Department of Economics, Second University of Naples, Italy, Corso Gran Priorato di Malta, Capua, e-mail: antonello.dambra@unina2.it

1 Introduction

Correspondence analysis is a widely used tool for obtaining a graphical representation of the interdependence between the rows and columns of a contingency table, and it is usually performed by applying a generalized singular value decomposition to the standardised residuals of a two-way contingency table obtaining a dimensionality reduction of the space. This decomposition ensures that the maximum information regarding the association between the two categorical variables are accounted for in a factorial plane of a correspondence plot, enabling us to visually understand the nature of this association. There are many ways in which the graphical display may be obtained and, usually, they use a dimension reduction of the global space. Dimension reduction can be achieved via a suite of interrelated methods for a two-way contingency table. We draw our attention to singular value decomposition (hereafter SVD), generalised singular value decomposition (GSVD), bivariate moment decomposition (BMD), and hybrid decomposition (HMD). Aim of this paper is to present an unified approach to several methods of correspondence analysis of a two way contingency table with ordinal categorical variables.

2 Basic notation

We consider samples from I different populations $(A_1,\dots,A_I)$, each of which is divided into J categories $(B_1,\dots,B_J)$. We assume that the samples of sizes $n_1,\dots,n_I$ from different populations are independent and that each sample follows a multinomial distribution. The probability of having an observation falls in the i-th row and j-th column of the table is denoted π_{ij}. An $I\times J$ contingency table $\mathbf{N}$ of the observations n_{ij} $(i=1,...,I; j=1,\dots,J)$ is considered with $n=\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}$. The $(i,\,j)$-th element of the relative frequencies matrix $\mathbf{P}$ is defined as $p_{ij}=n_{ij}/n$ such that $\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij}=1$.

Finally, let's suppose that $\mathbf{N}$ has one ordered set of categories (column) with row and column marginal frequencies given by $p_{i.}=\sum_{j=1}^{J} p_{ij}$ and $p_{.j}=\sum_{i=1}^{I} p_{ij}$, respectively. Moreover, let $\mathbf{D}_I$ and $\mathbf{D}_J$ represent the diagonal matrices of row and column marginal relative $p_{i.}$ and $p_{.j}$, respectively, with $\mathbf{r}=\mathbf{D}_I\mathbf{1}$ and $\mathbf{c}=\mathbf{1}^T\mathbf{D}_J$.

3 Correspondence analysis and its extensions based on SVD

3.1 Correspondence analysis

Correspondence analysis (hereafter CA) of cross-classifications between two categorical variables is usually presented as a multivariate method that decomposes the chi-squared statistic associated with a contingency table into orthogonal factors. Row and column categories are usually displayed in two-dimensional graphical form. This approach has been described can be described from other points of view. For instance, Goodman [14] shows that CA can be performed by applying the Singular Value Decomposition (SVD) on the Pearson's ratios table [14]. That is, for the $I \times J$ correspondence matrix $\mathbf{P}$ then its Pearson ratio α_{ij} is decomposed so that $\alpha_{ij} = \frac{p_{ij}}{p_{i.}p_{.j}} = 1 + \sum_{m=1}^{K} \lambda_m a_{im} b_{jm}$ with a_{im} and b_{jm} $\{m = 1, \ldots, K = \min(I, J) - 1\}$ singular vectors associated with the i'th row and j'th column category, respectively, and λ_m is the m'th singular value of the ratio. Moreover, these quantities are such to satisfy the conditions $\sum_j p_{.j} b_{jm} = \sum_i p_{i.} a_{im} = 0$ and $\sum_j p_{.j} b_{jm} b_{jm'} = \sum_i p_{i.} a_{im} a_{im'} = 1$ for $m = m'$, 0 otherwise.

Using the matrix notation, the above least squares estimates are obtained by a generalized singular value decomposition (GSVD) of matrix Ω

$$\Omega = \mathbf{D}_I^{-1}(\mathbf{P} - \mathbf{D}_I \mathbf{1}\mathbf{1}^T \mathbf{D}_J)\mathbf{D}_J^{-1} = \mathbf{A}\mathbf{D}_\lambda \mathbf{B}^T$$

with $\mathbf{A}^T \mathbf{D}_I \mathbf{A} = \mathbf{I}$, $\mathbf{B}^T \mathbf{D}_J \mathbf{B} = \mathbf{I}$ and where $\mathbf{D}_\lambda$ is a diagonal matrix with all singular values λ_m in descending order. It is well known that

$$\phi^2 = \frac{\chi^2}{n} = trace(\mathbf{D}_I^{-1/2}(\mathbf{P} - \mathbf{D}_I \mathbf{1}\mathbf{1}^T \mathbf{D}_J)\mathbf{D}_J^{-1}(\mathbf{P} - \mathbf{D}_I \mathbf{1}\mathbf{1}^T \mathbf{D}_J)^T \mathbf{D}_I^{-1/2}) = \sum_{m=1}^{K} \lambda_m^2$$

where χ^2 is the Pearson's chi squared statistic. See [5] for a bibliographic review, and for the new theoretical advances in this topic that have been made over the last 20 years.

3.2 Non symmetrical correspondence analysis

In two-way contingency tables, rows and columns often assume an asymmetric role. This aspect is not taken into account by correspondence analysis where it is supposed a symmetric role between the categorical variables with the decomposition of the Pearson's chi squared statistic. When the variables are asymmetrical related D'Ambra and Lauro [10] introduced a new approach named non symmetrical correspondence analysis which aim is to examine predictive relationships between rows and columns of a contingency table for which it is assumed that columns depend on rows, but not vice versa. This approach provides a visualisation of the magnitude of the measure of increase in predictability of a categorical response variable J given a categorical predictor variable I.

Non symmetrical correspondence analysis (NSCA) amounts to the following GSVD

$$\mathbf{D}_I^{-1}(\mathbf{P} - \mathbf{D}_I \mathbf{1}\mathbf{1}^T \mathbf{D}_J) = \mathbf{A}\mathbf{D}_\lambda \mathbf{B}^T$$

with $\mathbf{A}^T \mathbf{D}_I \mathbf{A} = \mathbf{I}$, and $\mathbf{B}^T \mathbf{B} = \mathbf{I}$. This GSVD leads to decompose the following quantity

$$N_\tau = n \times trace(\mathbf{D}_I^{-1/2}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)^T\mathbf{D}_I^{-1/2}) = n \times \sum_{m=1}^{K}\lambda_m^2$$

where N_τ is the numerator of the Goodman-Kruskal's tau index [15]

$$\tau = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} p_{i.}\left(\frac{p_{ij}}{p_{i.}} - p_{.j}\right)^2}{1-\sum_{j=1}^{J} p_{.j}^2} = \frac{N_\tau}{1-\sum_{j=1}^{J} p_{.j}^2}$$

that is a measure of predicability power of the the rows on columns. The conditional probability of j given row i is given by $p_{j|i} = p_{ij}/p_{i.}$ and the unconditional probability of column j by the marginal probability $p_{.j}$. The predictive power of row i on column j is thus computed by $p_{ij}/p_{i.} - p_{.j}$. When, for each column, the distribution of the response categories across each of the rows is identical to the overall marginal proportion there is no relative increase in predictability and thus τ is zero. Similarly, $\tau = 1$ only when there is perfect predictability of the response categories (columns) given the predictor categories (rows). A test criterion called C has been developed by Light and Margolin (1971) to assess the statistical significance of dependence. It is asymptotically approximated as a $\chi^2_{(I-1)(J-1)}$ random variable under the null hypothesis that the joint probabilities π_{ij} are equal to the column marginal probabilities $\pi_{.j}$, i.e. $H_0 : \pi_{ij} = \pi_{.j}$.

3.3 The decomposition of cumulative chi squared statistic (columns)

Beh, D'Ambra and Simonetti [3, 4] perform CA when one of the cross-classified variables has an ordered structure. They take into account the presence of an ordinal categorical variable by considering the cumulative sum of cell frequencies across the variable. Main aim of this approach is to determine graphically how similar cumulative categories are with respect to nominal ones. Let $z_{ik} = \sum_{j=1}^{k} n_{ij}$ be the cumulative frequency of the i-th row category up to the k-th column category providing a way of ensuring that the ordinal structure of the column categories is preserved. Similarly, let $d_k = \sum_{j=1}^{k} n_{.j}/n = \sum_{j=1}^{k} p_{.j}$ be the cumulative relative frequency up to the k-th column category. Moreover, let $\mathbf{W}$ be the $((J-1)\times(J-1))$ diagonal matrix of weights w_j and $\mathbf{M}$ a $((J-1)\times J)$ lower triangular matrix of ones. CA of cumulative frequencies (TA) amounts thus to the GSVD

$$\mathbf{D}_I^{-}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{M}^T\mathbf{W}^{\frac{1}{2}} = \mathbf{U}\mathbf{D}_\lambda\mathbf{V}^T$$

with $\mathbf{U}^T\mathbf{D}_I\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. It can be shown that

$$T = n \times trace(\mathbf{D}_I^{-\frac{1}{2}}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{M}^T\mathbf{W}\mathbf{M}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)^T\mathbf{D}_I^{-\frac{1}{2}}) = \sum_{i=1}^{I}\lambda_i^2$$

where $T = \sum_{j=1}^{J-1} w_j \left[\sum_{i=1}^{I} n_{i.}(\frac{z_{ij}}{n_{i.}} - d_j)^2\right]$ is the Taguchi's statistic [25, 26], with $0 \le T \le [n(I-1)]$ where w_j's are suitable weights. Possible choices for w_j could be to assign constant weights to each term ($w_j = 1/J$) or assume it proportional to the inverse of the conditional expectation of the k-th term under the hypothesis of independence ($w_j = 1/[d_j(1-d_j)]$). Taguchi's statistic is a measure of the association between categorical variables. It performs better Pearson chi-squared statistic when there is an order in the categories on the columns of the contingency table and it is more suitable in studies (such as clinical trials) where the number of categories within a variable is equal to, or larger than 5 [17]. The Taguchi's statistic was originally developed for the one-way Anova model for industrial experiments to test the hypothesis of homogeneity against monotonicity in the treatment effects. An $I \times J$ contingency table with row multinomial model with equal row totals ($n_{i.} = K$ observations per level of a factor A with I levels) has been then obtained. For this model, Nair [21] shows that the sum of squares for the factor A is given by $SSA = n\sum_{j=1}^{J-1}\sum_{i=1}^{I}(z_{ij}/K - d_j)^2/(d_j(1-d_j))$ which is the Taguchi's CSS statistic T with fixed and equal rows totals. Correspondence analysis of cumulative frequencies is then valid also for this special case of contingency table.

Nair [21] highlighted the link between the Pearson chi-squared statistic and the Taguchi's statistic: $T = \sum_{j=1}^{J-1}\chi_j^2$ where χ_j^2 is the Pearson chi-squared for the $I \times 2$ contingency tables obtained by aggregating the first j column categories and the remaining categories $(j+1)$ to J, respectively. For this property, Taguchi's statistic T has also been referred as the "Cumulative Chi-squared Statistic" (CSS) by Takeuchi and Hirotsu [27]. Moreover, Nair [21] showed that the distribution of T can be approximated using the Satterthwaites method [24].

We highlight that the above Taguchi statistic can be generalized to give the class of CSS-type tests given by $T_{\alpha_j} = \sum_{j=1}^{J-1}\alpha_j\chi_j^2$. If $\alpha_j = 1$ then $T_{\alpha_j} = T$, while T_{α_j} subsumes the simple alternative to the Taguchi CSS statistic proposed by Nair [21] with $\alpha_j = [d_j(1-d_j)]/J$. Beh et al. [3] show also that if $\mathbf{M}^T\mathbf{W}\mathbf{M} = \mathbf{I}$ then the Taguchi's statistic T amounts to the numerator of the Goodman-Kruskal τ index [15] and, for this reason, the Taguchi's statistic T can be viewed as the cumulative version of the τ index.

It is also interesting to highlight the relationship between the coordinates of the cumulative CA (CumCA) and those of ordinary CA. For cumulative CA the row coordinates are given by $\mathbf{O}_{CumCA} = (\mathbf{D}_I^{-1}\mathbf{P} - \mathbf{1}_I\mathbf{c}^T)\mathbf{M}^T\mathbf{W}^{1/2}\mathbf{V}$ and by $\mathbf{O}_{CA} = (\mathbf{D}_I^{-1}\mathbf{P} - \mathbf{1}_I\mathbf{c}^T)\mathbf{B}$ for ordinary CA, where $\mathbf{V}$ and $\mathbf{B}$ are the matrices containing the right singular vectors for cumulative CA and ordinary CA, respectively. It can be shown that cumulative CA coordinates can be directly computed from those of ordinary CA as $\mathbf{O}_{CumCA} = \mathbf{O}_{CA}\mathbf{B}^T\mathbf{M}^T\mathbf{W}^{1/2}\mathbf{V}$. In the same way we have $\mathbf{O}_{CA} = \mathbf{O}_{CumCA}\mathbf{V}^T(\mathbf{M}^T\mathbf{W}^{1/2})^{-1}\mathbf{B}$.

Nair [21, 22] showed also several properties of CSS test decomposing this statistics into orthogonal components by first providing it a matrix form. Let $\tilde{\mathbf{D}}$ be the $((J-1)\times J)$ matrix involving the cumulative column relative marginal frequencies d_j, respectively, with

$$\tilde{\mathbf{D}} = \begin{bmatrix} 1-d_1 & -d_1 & -d_1 & -d_1 \\ 1-d_2 & 1-d_2 & -d_2 & -d_2 \\ .... & ... & .... & .... \\ 1-d_{J-1} & 1-d_{J-1} & 1-d_{J-1} & -d_{J-1} \end{bmatrix}$$

In addition, let $\mathbf{M}$ be a $((J-1)\times J)$ lower triangular matrix of ones, and $\mathbf{y}_i = [n_{i1},\dots,n_{iJ}]^T$ be a vector of observed frequencies for the i-th row. Then, the CSS statistic can be written as $T = \sum_{i=1}^{I} \mathbf{y}_i^T \tilde{\mathbf{D}}^T \mathbf{W} \tilde{\mathbf{D}} \mathbf{y}_i / n_{i.} = trace(\mathbf{D}_I^{-1/2} \mathbf{N} \tilde{\mathbf{D}}^T \mathbf{W} \tilde{\mathbf{D}} \mathbf{N}^T \mathbf{D}_I^{-1/2})$. The matrix $\tilde{\mathbf{D}}^T \mathbf{W} \tilde{\mathbf{D}}$ could be then decomposed such that $\tilde{\mathbf{D}}^T \mathbf{W} \tilde{\mathbf{D}} = \mathbf{Q} \mathbf{D}_\lambda \mathbf{Q}^T$. Under the assumption of equiprobable columns ($w_j = J/[j\times(J-j)]^{-1}$ with $j = 1,\dots,J-1$), the eigenvectors are linked to the j-th degree Chebychev polynomial on the integers $\{1,\dots,J\}$. In this case the first (linear) and the second (quadratic) component are equivalent to the Wilcoxon statistic for the $2\times J$ table and to the Moods test [20], respectively.

Let $\pi_{ij}^c = \sum_{t=1}^{j} \pi_{it}$ be the cumulated probability of that an observation falls into the cumulated (i,j)-th cross-category, with $\pi_{.j}^c = \sum_{i=1}^{I} \pi_{ij}^c$ (common cumulative probability), and consider the null hypothesis $H_0 : \pi_{ij}^c = \pi_{.j}^c$. Denote with l_j the log-likelihood function (free model) of the j-th $I\times 2$ table obtained by aggregating the first j column categories and the remaining $(j+1)$ to J, and with $l_j^{H_0}$ the log-likelihood function of the same table under the null hypothesis (null model). Let's consider the following quantity $T_{\alpha_j}^{LR} = \sum_{j=1}^{J-1} \alpha_j LR_j$ with $LR_j = 2(l_j/l_J^{H_0})$ and where it is possible to show that LR_j is asymptotically conformable to a χ_j^2 using a finite number of terms of a Taylor expansion. If $\alpha_j = 1$ then $T_{\alpha_j}^{LR} = T$, while $T_{\alpha_j}^{LR}$ subsumes the simple alternative to the Taguchi CSS statistic proposed by Nair [21] with $\alpha_j = [d_j(1-d_j)]/J$.

3.4 The decomposition of cumulative chi squared statistic (rows and columns)

Taguchi's statistic has been developed to measure the association between categorical variables when there is an order in the categories on the columns of the contingency table. Beh, D'Ambra and Camminatiello [9] proposed a generalisation of the Taguchi decomposition based on cumulative frequencies for the rows and columns (HDA). This approach introduces two suitable cumulative matrices $\mathbf{R}$ and $\mathbf{C}$ to pool the rows and columns of contingency table.

Let $\mathbf{R}$ be a $2(I-1)\times I$ matrix formed by alternating the rows of an $(I-1)\times I$ lower triangular matrix of ones with the rows of an $(I-1)\times I$ upper triangular matrix of ones (by first removing the row consisting of all ones in both matrices). Similarly, $\mathbf{C}$ is a $J\times 2(J-1)$ matrix obtained by alternating the columns of an $J\times(J-1)$ upper triangular matrix of ones with the columns of an $J\times(J-1)$ lower triangular matrix of ones (by first removing the column consisting of all ones in both matrices). That is

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \text{ and } \mathbf{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & \dots & 1 & 0 \\ 0 & 1 & 1 & 0 & \dots & 1 & 0 \\ 0 & 1 & 0 & 1 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & 0 & 1 & \dots & 0 & 1 \end{bmatrix}.$$

Moreover, $\mathbf{P}_R$ and $\mathbf{P}_C$ ($\tilde{\mathbf{D}}_R$ and $\tilde{\mathbf{D}}_C$) are the diagonal matrices with the marginal relative (absolute) frequencies $h_{i.}$ and $h_{.j}$ of the doubly cumulative table $\mathbf{H} = \mathbf{RPC}$ ($\mathbf{H} = \mathbf{RNC}$), respectively.

The approach suggested by D'Ambra et al. [9] amounts to the SVD of the matrix $\mathbf{D}_R^{-\frac{1}{2}}\mathbf{R}(\mathbf{P} - \mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{C}^T\mathbf{D}_C^{-\frac{1}{2}}$ or to GSVD$[\tilde{\mathbf{D}}_R^{-}\mathbf{R}(\mathbf{P} - \mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{C}^T\tilde{\mathbf{D}}_C^{-}]_{\tilde{\mathbf{D}}_R,\tilde{\mathbf{D}}_C}$.

They show that this approach leads to consider the decomposition of Hirotsu's index $\chi^{**2} = \sum_{i=1}^{I-1}\sum_{j=1}^{J-1}\chi_{ij}^{**2}$ [17]

$$\begin{aligned}\chi^{**2} =& n(I-1)(J-1)trace(\mathbf{D}_R^{-\frac{1}{2}}\mathbf{R}(\mathbf{P} - \mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{C}^T\mathbf{D}_C^{-1}\mathbf{C}(\mathbf{P} - \mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)^T\mathbf{R}^T\mathbf{D}_R^{-\frac{1}{2}}) \\ =& n(I-1)(J-1)\sum_{k=1}^{K}\lambda_k^2.\end{aligned}$$

Hirotsu [17] introduced the doubly cumulative chi-squared statistic χ^{**2} in order to measure the association between two ordered categorical variables in a two-way contingency table, where χ_{ij}^{**2} is the chi-squared statistic for the 2×2 contingency table obtained by pooling the original table $I \times J$ data at the $i-$th row and j-th column. This index can be then viewed like a natural extension of Taguchi's statistic T for two ordered categorical variables. Hirotsu [18] showed also that the null distribution of the statistic χ^{**2} is approximated by $d\chi_v^2$ with $d = d_1 \times d_2$ and $v = (I-1)(J-1)/d$, where $d_1 = 1 + 2(J-1)^{-1}[\sum_{k=1}^{J-2}(\sum_{s=1}^{k}\lambda_s)/\lambda_{k+1}]$ e $d_2 = 1 + 2(I-1)^{-1}[\sum_{=1}^{I-2}(\sum_{s=1}^{k}\gamma_s)/\gamma_{k+1}]$ con $\lambda_s = (\sum_{h=1}^{s}n_{.h})/(\sum_{g=s+1}^{J}n_{.g})$ e $\gamma_s = (\sum_{h=1}^{s}n_{h.})/(\sum_{g=s+1}^{I}n_{g.})$. See [9] for deeper theoretical aspects.

Another approach to deal with the study of the association between ordered categorical variables has been suggested in literature by Cuadras and Cuadras [6] named CA based on double accumulative frequencies. The Cuadras and Cuadras's method (DA) is based on the SVD of the matrix $\mathbf{D}_I^{-\frac{1}{2}}\mathbf{L}(\mathbf{P} - \mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{M}^T\mathbf{W}^{\frac{1}{2}}$ which amounts to GSVD$[\mathbf{D}_I^{-}\mathbf{L}(\mathbf{P} - \mathbf{P}_I\mathbf{1}\mathbf{1}^T\mathbf{P}_J)\mathbf{M}^T\mathbf{W}^{\frac{1}{2}}]_{\mathbf{D}_I,\mathbf{I}}$ where $\mathbf{L}$ is a upper triangular matrix of ones. Main remark of this approach are that it does not seem to lead to the decomposition of any known association index, and matrices $\mathbf{L}$ and $\mathbf{M}$ pool the rows and the columns of table in a successive manner such that they do not provide the necessary $2(I-1) \times 2(J-1)$ 2×2 tables to compute Hirotsu's index χ^{**2}.

4 Correspondence analysis of ordinal cross-classifications based on the bivariate moment decomposition

Consider the rectangular matrix $\Omega = \mathbf{D}_I^{-1/2}(\mathbf{P} - \mathbf{D}_I \mathbf{1}\mathbf{1}^T \mathbf{D}_J)\mathbf{D}_J^{-1/2}$ of size $I \times J$ with $\mathbf{R} = \Omega\Omega^T$ and $\mathbf{C} = \Omega^T\Omega$. It is well known [5] that the generalised singular value decomposition of the matrix of Pearson contingencies yields results that are equivalent to those obtained by performing an eigen-decomposition on the matrices $\mathbf{R}$ and $\mathbf{C}$, as considered in Hills [16] description of reciprocal averaging. The Hill's reciprocal averaging procedure utilises the Gram–Schmidt orthogonalisation process for ensuring that the set of row and column scores, treated as the basis vectors, are orthogonally constrained with respect to $p_{i.}$ and $p_{.j}$. An alternative and reliable method of calculating the orthogonal vectors using the Gram–Schmidt process is to use the recurrence formulae of Emerson [13]. Emersons formulae require calculation of only the previous two polynomials and are especially well suited for the analysis of categorical variables that are ordinally structured. To reflect this ordinal structure, a set of ordered column scores $s_j(J)$ $(j = 1,\ldots,J)$ is usually used.

Using Emerson's orthogonal polynomials it is possible to decompose the total inertia ϕ^2 in different components, each of which represents a different power of the supposed relationship between row and column (linear, quadratic etc,). The advantage of using orthogonal polynomials relies in the fact that the power information is considered in the analysis and the resulting scoring scheme allows a clear interpretation of the linear, quadratic or higher order trend components. This kind of decomposition of the total inertia is at the core of several CA extensions [5].

4.1 Double Ordered Correspondence Analysis

Double Ordered Correspondence Analysis [1] (DOCA) decomposes the (i, j)-th Pearson ratio α_{ij} so that $\alpha_{ij} = 1 + \sum_{u=1}^{I-1}\sum_{v=1}^{J-1} a_u(i) b_v(j) z_{uv}$. For this method of decomposition, $z_{uv} = \sqrt{n}\sum_{i,j} a_u(i)\, b_v(j)\, p_{ij}$ is the (u,v)th generalised correlation [12] where $\{a_u(i) : u = 1,\ldots,I-1\}$ and $\{b_v(j) : v = 1,\ldots,J-1\}$ are the orthogonal polynomials [13] for the i-th row and j-th column respectively. The bivariate association z_{uv} are collected in $\mathbf{Z} = \mathbf{A}_*^T\mathbf{P}\mathbf{B}_*$ where $\mathbf{A}_*$ contains the $I-1$ non-trivial row orthogonal polynomials and $\mathbf{B}_*$ is the $J \times (J-1)$ matrix of the $J-1$ non-trivial column orthogonal polynomials. The matrix Ω can be then rewritten as

$$\mathbf{D}_I^{-1/2}(\mathbf{P} - \mathbf{D}_I \mathbf{1}\mathbf{1}^T \mathbf{D}_J)\mathbf{D}_J^{-1/2} = \mathbf{A}_*\mathbf{Z}\mathbf{B}_*^T$$

with $\mathbf{A}_*^T\mathbf{D}_I\mathbf{A}_* = \mathbf{I}$ and $\mathbf{B}_*^T\mathbf{D}_J\mathbf{B}_* = \mathbf{I}$. This kind of decomposition of matrix Ω has been named "Bivariate Moment Decomposition".

It is possible to show [1] that the elements of $\mathbf{Z}$ (that is, the bivariate associations z_{uv}) are asymptotically standard normal and indipendent. Moreover, Rayner and Best [23] showed that the Pearson chi-squared statistic can be decomposed into the sum of squares of the generalized correlations so that $\phi^2 = \chi^2/n = \sum_{u=1}^{I-1}\sum_{v=1}^{J-1} z_{uv}^2$ and sources of variation for the row and column profiles can be obtained. Observe that the above chi-squared index is partitioned into $(I-1)(J-1)$ terms, where the significance of each term can be compared with the χ^2 with one degree of freedom (dof). Sources of variation for the row and column profiles can be easily obtained. For instance, any difference in the row profiles in terms of their location is computed by $\sum_{v=1}^{J-1} z_{1v}^2$ while the row dispersion component is given by $\sum_{v=1}^{J-1} z_{2v}^2$ .The significance of each component can be compared with the χ^2 with $(J-1)$ dof. Similarly, location and dispersion column components can be computed.

This approach to correspondence analysis uses then the bivariate moment decomposition to identify linear (location), quadratic (dispersion) and higher order moments. Note that this feature is not readily available by using classical SVD.

4.2 Double ordered non symmetric correspondence analysis

For the analysis of the cross-classification of two ordinal variables, Lombardo, Beh and D'Ambra [19] propose the BMD decomposition of the Goodman and Kruskal index and the graphical visualization of the data association structure for a two-way data matrix. The methodology named doubly ordinal non symmetric correspondence analysis (DONSCA) is designed to allow the user to visualize the dependence relationship between categories of a response and a predictor variable in terms of components that reflect sources of variation in terms of the location (mean), dispersion (spread) and higher order moments. Rather than considering the singular value decomposition of the matrix $\mathbf{D}_I^{-1}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)$ after computing orthonormal polynomials, an alternative strategy is to apply bivariate moment decomposition to the centred row profiles such that $\mathbf{D}_I^{-1}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J) = \mathbf{A}_*\Lambda\mathbf{B}_*$ where $\Lambda = \mathbf{A}_*^T\mathbf{D}_I\mathbf{P}\mathbf{B}_*$ with $\mathbf{A}_*^T\mathbf{D}_I\mathbf{A}_* = \mathbf{I}_{I-1}$ and $\mathbf{B}_*^T\mathbf{B}_* = \mathbf{I}_{J-1}$. Lombardo et al. [19] showed that the numerator of the Goodman–Kruskal τ index can be decomposed into the sum of squares of the generalized correlations so that $N_\tau = \sum_{u=1}^{I-1}\sum_{v=1}^{J-1}\lambda_{uv}^2$. Each term λ_{uv}^2 shows the quality of the symmetric/asymmetric association of the ordered categorical variables. For example, the linear component for the row variable is given by $\lambda_{1.}^2$, while the relevance of the linear component for the column variable is showed by can be found by $\lambda_{.1}^2$. The significance overall predicability can be tested by the C statistics given by $C = \sum_{u=1}^{I-1}\sum_{v=1}^{J-1}\tilde{\lambda}_{uv}^2$ where each term $\tilde{\lambda}_{uv}^2 = \lambda_{uv}^2[(n-1)(J-1)/(1-\sum_{j=1}^{J} p_{.j}^2)]^{1/2}$ is a random variable from an asymptotically standard normal distribution [8, 11] and $C \sim \chi^2_{(I-1)(J-1)}$ and each $\tilde{\lambda}_{i.}^2 \sim \chi^2_{(J-1)}$.

5 Correspondence analysis based on the hybrid moment decomposition

5.1 Singly ordered correspondence analysis

An alternative approach to partitioning the Pearson chi-squared statistic for a two-way contingency table with one ordered set of categories is given by the Singly Ordered Correspondence Analysis [1] (SOCA). This method combines the approach of orthogonal polynomials for the ordered columns and singular vectors for the unordered rows (named hybrid moment decomposition), such that $\chi^2 = \sum_{u=1}^{M^*}\sum_{v=1}^{J-1}\mathbf{Z}_{(u)v}^2$ with $M^* \leq I-1$ and where $z_{(u)v} = \sqrt{n}\sum_{i,j} p_{ij}a_{iu}b_v(j)$ are asymptotically standard normally distributed random variables. The parentheses around u indicates that the above formulas are concerned with a non-ordered set of row categories. Quantities $z_{(u)v}$ can be written in matrix notation as $\mathbf{Z} = \mathbf{A}^T\mathbf{P}\mathbf{B}_*$ where $\mathbf{A}$ is the $I \times (I-1)$ matrix of left singular vectors. 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. The Pearson ratio is given by $\alpha_{ij} = \sum_{u=0}^{M^*}\sum_{v=0}^{J-1} a_{iu}(z_{(u)v}/\sqrt{n})b_v(j)$. Eliminating the trivial solution, the matrix Ω can be also rewritten as

$$\mathbf{D}_I^{-1/2}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{D}_J^{-1/2} = \mathbf{A}\mathbf{Z}\mathbf{B}_*^T$$

with $\mathbf{A}^T\mathbf{D}_I\mathbf{A} = \mathbf{I}$ and $\mathbf{B}_*^T\mathbf{D}_J\mathbf{B}_* = \mathbf{I}$. For deeper information refer to [1] [2].

5.2 Singly ordered non symmetric correspondence analysis

When a two-way contingency table consists of only one ordered variable, a hybrid decomposition using a combination of singular vectors for the nominal variable and orthogonal polynomials for the ordered variable can be applied. Lombardo, Beh and D'Ambra [19] suggested a methodology, named singly ordinal non-symmetric correspondence analysis, that allows to combine the summaries obtained from SVD and BMD of the data. It allows to visualize and identify the primary causes of the dependent relationship between the categories of two variables where one of them is assumed to be an ordinal categorical variable. The total inertia as well as the partial inertia can be expressed by components that reflect within- and between-variable variation in terms of location, dispersion and higher-order moments. For this approach, the numerator of the Goodman–Kruskal tau index N_τ, is partitioned using generalised correlations. Authors illustrate two distinct approaches: one when the predictor variable consists of ordered categories (SONSCA1) and another when the response variable consists of ordered categories (SONSCA2).

In SONSCA1 hybrid decomposition uses orthogonal polynomials for the ordinal row categories and singular vectors for the nominal column categories such that

$$\mathbf{D}_I^{-1}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)=\mathbf{A}_*\Lambda\mathbf{B}$$

where $\Lambda=\mathbf{A}_*^T\mathbf{D}_I\mathbf{P}\mathbf{B}$ with $\mathbf{A}_*^T\mathbf{D}_I\mathbf{A}_*=\mathbf{I}_{I-1}$ and $\mathbf{B}^T\mathbf{B}=\mathbf{I}$. Lombardo et al. [?] showed that N_τ is decomposed as $N_\tau=\sum_{u=1}^{I-1}\sum_{m=1}^{M}\lambda_{um}^2$. The significance overall predicability can be tested by the C statistics given by $C=\sum_{u=1}^{I-1}\sum_{m=1}^{M}\tilde{\lambda}_{um}^2$ where each term $\tilde{\lambda}_{um}^2=\lambda_{um}^2[(n-1)(J-1)/(1-\sum_{j=1}^{J}p_{.j}^2)]^{1/2}$ is a random variable from an asymptotically standard normal distribution.

Similarly, In SONSCA2 hybrid decomposition uses instead singular vectors for the nominal row categories and orthogonal polynomials for the ordinal column categories such that

$$\mathbf{D}_I^{-1}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)=\mathbf{A}\Lambda\mathbf{B}_*$$

where $\Lambda=\mathbf{A}^T\mathbf{D}_I\mathbf{P}\mathbf{B}_*$ with $\mathbf{A}^T\mathbf{D}_I\mathbf{A}=\mathbf{I}$ and $\mathbf{B}_*^T\mathbf{B}_*=\mathbf{I}_{J-1}$. Lombardo et al. [19] showed that N_τ is decomposed as $N_\tau=\sum_{m=1}^{M}\sum_{v=1}^{J-1}\lambda_{mv}^2$. The significance overall predicability can be tested by the C statistics given by $C=\sum_{m=1}^{M}\sum_{v=1}^{J-1}\tilde{\lambda}_{mv}^2$ where each term $\tilde{\lambda}_{mv}^2=\lambda_{mv}^2[(n-1)(J-1)/(1-\sum_{j=1}^{J}p_{.j}^2)]^{1/2}$ is a random variable from an asymptotically standard normal distribution.

5.3 Hybrid cumulative correspondence analysis

D'Ambra [7] extends the properties of orthogonal polynomials to the Cumulative Correspondence Analysis with the aim to identify the linear, quadratic and higher order components of rows with respect to the aggregate columns categories. The joint effect of these methods leads to the decomposition of the Taguchi's index in power components. In hybrid cumulative correspondence analysis (HCCA) both variables present an ordinal structure and only for the response one (column) it has been considered the cumulative sum of cell frequencies across this variable.

In HCCA hybrid decomposition uses orthogonal polynomials for the ordinal row categories and singular vectors for the cumulative column categories such that

$$\mathbf{D}_I^{-1}(\mathbf{P}-\mathbf{D}_I\mathbf{1}\mathbf{1}^T\mathbf{D}_J)\mathbf{M}^T\mathbf{W}^{\frac{1}{2}}=\mathbf{A}_*\Lambda\mathbf{B}$$

where $\Lambda=\mathbf{A}_*^T\mathbf{D}_I\mathbf{P}\mathbf{B}_*$ with $\mathbf{A}_*^T\mathbf{D}_I\mathbf{A}_*=\mathbf{I}_{I-1}$ and $\mathbf{B}^T\mathbf{B}=\mathbf{I}$. D'Ambra [7] shows that Taguchi's index T is decomposed as $T=n\sum_{u=1}^{I-1}\sum_{m=1}^{M}\lambda_{um}^2$. Moreover, let $\mathbf{P}_j$ be the $I\times 2$ contingency table obtained by aggregating the first j column categories and the remaining categories $(j+1)$ to J of table $\mathbf{P}$. It can be shown that the linear, quadratic and higher order components of T can be written as sum of the correspondence components of rows computed for each matrix $\mathbf{P}_j$, respectively. For example, the linear components of T_L can be decomposed according to the sum of the linear components ${}^j\chi_L^2$ of matrices $\mathbf{P}_j$ with $j=1,\ldots,J-1$: $T_L=n\sum_{v=1}^{J-1}z_{1v}^2=\ {}^1\chi_L^2+\ {}^2\chi_L^2+\cdots+\ {}^{J-1}\chi_L^2$.

Similarly, the quadratic and higher order components of T can be obtained. The T statistics can be also written as $T = \sum_{j=1}^{J-1}\sum_{i=1}^{I-1} {}^{j}\tilde{\lambda}_i^2$ where each element ${}^{j}\tilde{\lambda}_i$ is a statistical variable that follows an asymptotically standard normally distribution. Moreover, each ${}^{j}\tilde{\lambda}_i^2$ follows a chi-square distribution with 1 degree of freedom.

6 A unified approach

All these approaches can be unified in a single theoretical framework. Let's consider the following factorization of a matrix into a product of matrices that we name "generalized factorization of a matrix" (GFM).

Let Γ and Φ be given positive definite symmetric matrices of order $(n \times n)$ and $(p \times p)$, respectively. The GFM of matrix $\mathbf{X}$ is defined as $\mathbf{X} = \mathbf{U}\Lambda\mathbf{V}^T$ where the columns of $\mathbf{U}$ and $\mathbf{V}$ are orthonormalized with respect to Γ and Φ (that is $\mathbf{U}^T\Gamma\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\Phi\mathbf{V} = \mathbf{I}$), respectively, and Λ is a positive definite matrix. It is noted as $\text{GFM}(\mathbf{X})_{\Omega,\Phi}$. GFM can take several forms. For instance, if the used GFM is the GSVD then matrices $\mathbf{U}$ and $\mathbf{V}$ are given by $\mathbf{A}$ and $\mathbf{B}$ that are orthonormalized with respect to Γ and Φ (that is $\mathbf{A}^T\Gamma\mathbf{A} = \mathbf{I}$ and $\mathbf{B}^T\Phi\mathbf{B} = \mathbf{I}$), respectively, with $\Lambda = \mathbf{D}_\lambda$ is a diagonal and positive definite matrix containing the generalized singular values, ordered from largest to smallest. They can be obtained by means the ordinary SVD. Similarly, if the used GFM is instead the BMD then matrices $\mathbf{U}$ and $\mathbf{V}$ are given by the the vectors of $\mathbf{A}_*$ and $\mathbf{B}_*$ orthonormalized with respect to Γ and Φ (that is $\mathbf{A}_*^T\Gamma\mathbf{A}_* = \mathbf{I}$ and $\mathbf{B}_*^T\Phi\mathbf{B}_* = \mathbf{I}$), respectively, with $\Lambda = \mathbf{A}_*^T\mathbf{P}\mathbf{B}_*$ (Λ in this case is not a diagonal matrix).

In order to represent the rows and columns of $\mathbf{N}$ we can consider the following unifying expression

$$\text{GFM}[\mathbf{D}_{(R)}^{-}\hat{\mathbf{R}}(\mathbf{P} - \mathbf{P}_I\mathbf{1}_I\mathbf{1}_J^T\mathbf{P}_J)\hat{\mathbf{C}}^T\mathbf{W}^{\frac{1}{2}}\mathbf{D}_{(C)}^{-}]_{\mathbf{D}_{(R)},\mathbf{D}_{(C)}} \tag{1}$$

that is equivalent to write $\mathbf{D}_{(R)}^{-1/2}\hat{\mathbf{R}}(\mathbf{P} - \mathbf{P}_I\mathbf{1}_I\mathbf{1}_J^T\mathbf{P}_J)\hat{\mathbf{C}}^T\mathbf{W}^{\frac{1}{2}}\mathbf{D}_{(C)}^{-1/2} = \mathbf{U}\Lambda\mathbf{V}^T$.

According to type of factorization used (GSVD/SVD, BMD and HMD), this GFM let us to subsume all the previous approaches according to the structure of the matrices concerned (see Table 1). Methods for which the ordinal structure of a categorical variable is directly taken into account in their formulations and calculation are listed in bold character in Table 1.

For instance, if $\mathbf{D}_{(R)} = \mathbf{D}_I$, $\hat{\mathbf{R}} = \mathbf{I}$, $\hat{\mathbf{C}} = \mathbf{I}$, $\mathbf{W} = \mathbf{I}$ and $\mathbf{D}_{(C)} = \mathbf{D}_J$, with $\mathbf{U} = \mathbf{A}$, $\mathbf{V} = \mathbf{B}$ and $\Lambda = \mathbf{D}_\lambda$, then $\text{GFM}[\mathbf{D}_{(R)}^{-}\hat{\mathbf{R}}(\mathbf{P} - \mathbf{P}_I\mathbf{1}_I\mathbf{1}_J^T\mathbf{P}_J)\hat{\mathbf{C}}^T\mathbf{W}^{\frac{1}{2}}\mathbf{D}_{(C)}^{-}]_{\mathbf{D}_{(R)},\mathbf{D}_{(C)}}$ amounts to ordinary Correspondence Analysis (CA). Double Ordered Correspondence Analysis (DOCA) is instead obtained by using in (1) $\mathbf{D}_{(R)} = \mathbf{D}_I$, $\hat{\mathbf{R}} = \mathbf{I}$, $\hat{\mathbf{C}} = \mathbf{I}$, $\mathbf{W} = \mathbf{I}$ and $\mathbf{D}_{(C)} = \mathbf{D}_J$, with $\mathbf{U} = \mathbf{A}_*$, $\mathbf{V} = \mathbf{B}$ and $\Lambda = \mathbf{A}_*^T\mathbf{P}\mathbf{B}_*$.

In the same way, if $\mathbf{D}_{(R)} = \mathbf{D}_I$, $\hat{\mathbf{R}} = \mathbf{I}$, $\hat{\mathbf{C}} = \mathbf{M}$, $\mathbf{W} = \mathbf{W}$ and $\mathbf{D}_{(C)} = \mathbf{I}_{J-1}$, with $\mathbf{U} = \mathbf{A}_*$, $\mathbf{V} = \mathbf{B}$ and $\Lambda = \mathbf{A}_*^T \mathbf{D}_I \mathbf{P} \mathbf{B}$, then $\text{GFM}[\mathbf{D}_{(R)}^{-} \hat{\mathbf{R}} (\mathbf{P} - \mathbf{P}_I \mathbf{1}_I \mathbf{1}_J^T \mathbf{P}_J) \hat{\mathbf{C}}^T \mathbf{W}^{\frac{1}{2}} \mathbf{D}_{(C)}^{-}]_{\mathbf{D}_{(R)}, \mathbf{D}_{(C)}}$ amounts to Hybrid Cumulative Correspondence Analysis (HCCA).

Table 1 A unified framework of CA with ordinal categorical data.

	$\text{GFM}[\mathbf{D}_{(R)}^{-} \hat{\mathbf{R}} (\mathbf{P} - \mathbf{D}_I \mathbf{1}_I \mathbf{1}_J^T \mathbf{D}_J) \hat{\mathbf{C}}^T \mathbf{W}^{\frac{1}{2}} \mathbf{D}_{(C)}^{-}]_{\mathbf{D}_{(R)}, \mathbf{D}_{(C)}}$											
			$\mathbf{D}_{(R)}^{-1/2} \hat{\mathbf{R}} (\mathbf{P} - \mathbf{D}_I \mathbf{1}_I \mathbf{1}_J^T \mathbf{D}_J) \hat{\mathbf{C}}^T \mathbf{W}^{\frac{1}{2}} \mathbf{D}_{(C)}^{-1/2} = \mathbf{U} \Lambda \mathbf{V}^T$									
Method	Row	Column	$\mathbf{D}_{(R)}$	$\hat{\mathbf{R}}$	$\hat{\mathbf{C}}$	$\mathbf{W}$	$\mathbf{D}_{(C)}$	$\mathbf{U}$	Λ	$\mathbf{V}$	GFM	Statistic
CA	Nom-Ord	Nom-Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{I}_J$	$\mathbf{D}_J$	$\mathbf{A}$	$\mathbf{D}_\lambda$	$\mathbf{B}$	SVD	ϕ^2
NSCA	Nom-Ord	Nom-Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{I}_J$	$\mathbf{I}_J$	$\mathbf{A}$	$\mathbf{D}_\lambda$	$\mathbf{B}$	SVD	N_τ
TA	Nom	Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{M}$	$\mathbf{W}$	$\mathbf{I}_{J-1}$	$\mathbf{A}$	$\mathbf{D}_\lambda$	$\mathbf{B}$	SVD	T
HDA	Ord	Ord	$\mathbf{D}_R$	$\mathbf{R}$	$\mathbf{C}^T$	$\mathbf{I}_{2\times(J-1)}$	$\mathbf{D}_C$	$\mathbf{A}$	$\mathbf{D}_\lambda$	$\mathbf{B}$	SVD	χ^{**2}
DA	Ord	Ord	$\mathbf{D}_I$	$\mathbf{L}$	$\mathbf{M}$	$\mathbf{W}$	$\mathbf{I}_{J-1}$	$\mathbf{A}$	$\mathbf{D}_\lambda$	$\mathbf{B}$	SVD	–
DOCA	Ord	Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{D}_J$	$\mathbf{A}_*$	$\mathbf{A}_*^T \mathbf{P} \mathbf{B}_*$	$\mathbf{B}_*$	BMD	ϕ^2
DONSCA	Ord	Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{I}_J$	$\mathbf{A}_*$	$\mathbf{A}_*^T \mathbf{D}_I \mathbf{P} \mathbf{B}_*$	$\mathbf{B}_*$	BMD	N_τ
SOCA	Nom	Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{D}_J$	$\mathbf{A}$	$\mathbf{A}^T \mathbf{P} \mathbf{B}_*$	$\mathbf{B}_*$	HMD	ϕ^2
SONSCA1	Ord	Nom	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{I}_J$	$\mathbf{A}_*$	$\mathbf{A}_*^T \mathbf{D}_I \mathbf{P} \mathbf{B}$	$\mathbf{B}$	HMD	N_τ
SONSCA2	Nom	Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{I}_I$	$\mathbf{I}_J$	$\mathbf{I}_J$	$\mathbf{A}$	$\mathbf{A}^T \mathbf{D}_I \mathbf{P} \mathbf{B}_*$	$\mathbf{B}_*$	HMD	N_τ
HCCA	Ord	Ord	$\mathbf{D}_I$	$\mathbf{I}_I$	$\mathbf{M}$	$\mathbf{W}$	$\mathbf{I}_{J-1}$	$\mathbf{A}_*$	$\mathbf{A}_*^T \mathbf{D}_I \mathbf{P} \mathbf{B}$	$\mathbf{B}$	HMD	T

where "Nom" and "Ord" stand for Nominal and Ordinal variable, respectively.

References

1. Beh, E.J.: Simple correspondence analysis of ordinal cross-classifications using orthogonal polynomials. Biometrical Journal, **39**, 589–613, 1997
2. Beh, E.J.: Simple Correspondence Analysis of Nominal-Ordinal Contingency Tables. J. Appl. Math. Decis. Sci., Article ID **218140**, doi:10.1155/2008/218140, (2008)
3. Beh, E.J., D'Ambra, L., Simonetti, B.: Ordinal correspondence analysis based on cumulative chi-squared test. In: Correspondence Analysis and Related Methods, Rotterdam: CARME 2007 Proceedings (2007)
4. Beh, E.J., D'Ambra, L., Simonetti, B.: Correspondence Analysis of Cumulative Frequencies using a Decomposition of Taguchi's Statistic. Commun. Stat. Theory Methods, **40**, 1620–1632, (2011)
5. Beh, E.J., Lombardo, R.: Correspondence Analysis: Theory, Practice and New Strategies, Wiley, (2014)
6. Cuadras, C.M., Cuadras, D.: A Unified Approach for the Multivariate Analysis of Contingency Tables. Open journal of statistics, **5**, 223–232, (2015)
7. D'Ambra, A.: Cumulative Correspondence Analysis using Orthogonal Polynomials, Commun. Stat. Theory. Methods., to appear.

8. D'Ambra, L., Beh, E.J., Amenta, P.: CATANOVA for two-way contingency tables with an ordinal response using orthogonal polynomials. Commun. Stat. Theory. Methods., **34**, 1755–1769, (2005).
9. D'Ambra, L., Beh, E.J., Camminatiello, I.: Cumulative Correspondence Analysis of Two-Way Ordinal Contingency Tables. Commun. Stat. Theory Methods, **43**, 6 1099–1113, (2014)
10. D'Ambra, L., Lauro, N.: Non symmetrical analysis of three-way contingency tables. In: Coppi, R., Bolasco, S. (eds.) Multiway Data Analysis, pp. 301-315. Amsterdam: Elsevier Science Publishers B.V., (1989)
11. DAmbra, L., Lombardo, R., Amenta, P.: Non symmetric correspondence analysis for ordered two-way contingency table. In: Proceedings of the XLI Italian Statistics Society Conference, University of Milano Bicocca, pp. 191–201, (2002)
12. Davy, P.J., Rayner, J.C.W., Beh, E.J.: Generalised correlations and Simpson's Paradox. In: Pemajayantha, V., Mellor, R.W., Peiris, S., Rajasekera, J.R. (eds.) Current Research in Modelling, Data Mining and Quantitative Techniques, pp. 63–73. University of Western Sydney, Sydney, Australia, (2003)
13. Emerson, P.L.: Numerical construction of orthogonal polynomials from a general recurrence formula. Biometrics, **24**, 696–701, (1968)
14. Goodman, L.A.: A single general method for the analysis of cross-classified data: Reconciliation and synthesis of some methods of Pearson, Yule, and Fisher, and also some methods of correspondence analysis and association analysis. J. Amer. Statist. Assoc. **91**, 408–428, (1996)
15. Goodman, L.A., Kruskal, W. H.: Measures of association for cross-classifications, J. Amer. Statist. Assoc. **49**, 732–764, (1954)
16. Hill, M.O.: Correspondence Analysis: a neglected multivariate method. J. R. Stat. Soc. Ser. C Appl. Stat., **23**, 340–354, (1974).
17. Hirotsu, C.: Cumulative Chi-squared Statistic as a Tool for Testing Goodness of Fit. Biometrika, **73**, 165–173, (1986)
18. Hirotsu, C.: Modelling and analysing the generalized interaction. Proceedings of the Third IEEE Conference on Control Applications, ISBN-10: 0780318722 (1994).
19. Lombardo, R., Beh, E.J., DAmbra, L.: Non-symmetric correspondence analysis with ordinal variables. Comput. Stat. Data Anal., **52**, 566–577, (2007)
20. Mood, A.M.: On the Asymptotic Efficiency of Certain Non-parametric Two-Sample Tests. Ann. of Math. Stat., **25**, 514–522, (1954)
21. Nair, V.N.: Testing in industrial experiments with ordered categorical data. Technometrics, **28**, 4, 283–291, (1986)
22. Nair, V.N.: Chi-squared type tests for ordered alternatives in contingency tables. J. Amer. Statist. Assoc. **82**, 283–291, (1987)
23. Rayner, J.C.W., Best, D.J.: Smooth extensions of Pearson's product moment correlation and Spearman's rho. Stat. Probabil. Lett., **30**, 171–177, (1996)
24. Satterthwaite, F.E.: An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin, **2**, 6, 110–114, (1946)
25. Taguchi, G.: Statistical Analysis. Tokyo: Maruzen, (1966)
26. Taguchi, G.: A new statistical analysis for clinical data, the accumulating analysis, in contrast with the chi-square test. Saishin Igaku, **29**, 806–813, (1974)
27. Takeuchi, K., Hirotsu, C.: The Cumulative Chi Square Method Against Ordered Alternative in Two-way Contingency Tables. Reports of Statistical Application Research. Japanese Union of Scientists and Engineers, **29**, 1–13, (1982)