CrossMark

An algebraic generalisation of some variants of simple correspondence analysis

Eric J. Beh[1] · **Rosaria Lombardo**[2]

Abstract For an analysis of the association between two categorical variables that are cross-classified to form a contingency table, graphical procedures have been central to this analysis. In particular, correspondence analysis has grown to be a popular method for obtaining such a summary and there is a great variety of different approaches that one may consider to perform. In this paper, we shall introduce a simple algebraic generalisation of some of the more common approaches to obtaining a graphical summary of association, where these approaches are akin to the correspondence analysis of a two-way contingency table. Specific cases of the generalised procedure include the classical and non-symmetrical correspondence plots and the symmetrical and isometric biplots.

Keywords Adjusted residual · Biplot · Correspondence analysis · Isometric plot · Non-symmetrical correspondence analysis · Singular Value Decomposition · Standardised Residual

✉ Eric J. Beh
eric.beh@newcastle.edu.au

Rosaria Lombardo
rosaria.lombardo@unicampania.it

[1] School of Mathematical and Physical Sciences, University of Newcastle, Callaghan, Australia

[2] Department of Economics, University of Campania, Caserta, Italy

1 Introduction

When analysing the association between two categorical variables that form a contingency table, there are typically two important aspects that one should consider. Firstly, the analyst can identify the strength of the association between the variables. This may be conducted by considering standard statistical tests such as Pearson's chi-squared test of independence. Secondly, where a statistically significant association is detected, one may determine those categories within and between the variables that are similar or different in terms of the conditional distribution of the cell counts. Graphical summaries provide a very simple and intuitive way to view these two characteristics. However different types of graphical summaries used to visualise this association (or lack of it) can lead to different conclusions about the way in which the variables are related.

There are many relatively sophisticated ways to visualise the association in two-way contingency tables. Some of the more popular approaches include using various types of biplots. One may consider, for example, the relatively recent extensive discussions of biplots by Gower et al. (2011) and Greenacre (2010a). One may also construct plots from any one of the many members of the correspondence analysis family; see Beh and Lombardo (2012, 2014) and Nishisato (2007) for an extensive and diverse list of such members.

In this paper, a simple algebraic generalisation is proposed that incorporates, as special cases, many aspects of correspondence analysis. Specifically, the generalisation involves decomposing the matrix of standardised residuals, but can also be extended to decompose other residuals and to incorporate other measures of association including the tau index of Goodman and Kruskal (1954) and the Freeman–Tukey statistic of Freeman and Tukey (1950). The most popular method of decomposition used is singular value decomposition (SVD) and it is this method of decomposition that is considered in this paper. However, there are others that can be incorporated into the proposed algebraic generalisation including the bivariate moment decomposition (BMD) used for graphically depicting the association between ordinal categorical variables; see Beh (1997, 1998), Best (1995), Best et al. (1999, 2000), Lombardo et al. (2007, 2016). Another type of decomposition is the hybrid of SVD and BMD discussed by Beh (2001) which Beh (2008) used to graphically analyse nominal-ordinal two-way contingency tables.

Irrespective of the method of decomposition used, the purpose of decomposing the residuals is to provide a mechanism for selecting a sub-set of dimensions from the optimal number required to visualise the relationship between the variables. This is achieved while maximising the association structure between the variables. Generally one, two or at most three dimensions are enough for a simple visualisation of the association structure. For a two-dimensional view of multi-dimensional data, many graphical procedures can be used, including Andrew's Curve Andrews (1972), Chernoff's face Chernoff (1973), cobweb diagrams (Upton 2000, 2016) and the mosaic display (Friendly 1994, 1999). One may also consider methods of constructing elliptical regions for simple correspondence analysis that can include information contained in dimensions higher than the second; see Beh (2010). Alternatively, bootstrap tech-

niques can be used for constructing such regions (Linting et al. 2007; Lombardo and Ringrose 2012; Markus 1994; Ringrose 2012).

We must point out that discussing generalisatons in the context of correspondence analysis are not new. For example, one may consider Choulakian (1988), De Leeuw (1993), Lorenzo-Seva et al. (2009), Velden and Kiers (2005) and Yanai (1986) as examples of such generalisations. What makes these contributions to the correspondence analysis literature different from what is described in this paper is that, in the past, generalisations have focused on showing how viewing correspondence analysis from a variety of perspectives leads to similar row/column scoring solutions. Our approach provides a broad, yet relatively simple, algebraic generalisation that incorporates, as special cases, the growing diversity of mathematical and graphical perspectives that have been developed over the past few decades.

To derive the simple algebraic generalisation of visualising categorical variables, this paper is divided into seven further sections. Section 2 defines the notation used throughout the paper and describes the application of SVD on the matrix of standardised residuals, yielding the classic approach to correspondence analysis. An unbalanced generalised framework for the analysis of two-way contingency tables is given in Sect. 3 and special cases of these are described in Sect. 4. A balanced generalised framework is discussed in Sect. 5 and special cases of it are looked at in Sect. 6. Section 8 provides some further insight into how these generalisations may be implemented. Some concluding remarks are made in Sect. 9.

2 Simple correspondence analysis

To begin our algebraic generalisation, we shall first briefly highlight some of the key features of the correspondence analysis of a two-way contingency table. There are a number of variations of correspondence analysis that can be considered, and our generalisation incorporates some of the commonly used approaches.

For a $r \times c$ two-way contingency table $\mathbf{N}$, denote its grand total by n. Also, denote the matrix of joint proportions by $\mathbf{P}$ so that $\mathbf{P} = \mathbf{N}/n$. Let $\mathbf{r}$ be the vector of row marginal proportions and $\mathbf{c}$ be the vector of column marginal proportions. Denote $\mathbf{D}_r$ as the diagonal matrix of row marginal proportions such that $\mathbf{D}_r = \text{diag}(\mathbf{r})$. Similarly, let $\mathbf{D}_c = \text{diag}(\mathbf{c})$.

Consider the decomposition of Pearson's matrix of standardised residuals

$$\mathbf{Z} = \mathbf{D}_r^{-1/2} \left(\mathbf{P} - \mathbf{rc}' \right) \mathbf{D}_c^{-1/2} = D^* \left(\mathbf{D}_\lambda, \ \mathbf{U}, \ \mathbf{V} \right). \tag{1}$$

The right hand side of (1), $D^*(\bullet)$, denotes the type of decomposition used. For this paper we will focus on the case where $D^*(\bullet)$ is the SVD of $\mathbf{Z}$, however other decomposition methods may be considered. Two such alternatives are the Bivariate Moment Decomposition (BMD) commonly used for two ordinal variable or the Hybrid Decomposition (HD) for a singly ordered two-way contingency table that combines specific features of SVD and BMD into a single decomposition. When the standardised residuals are considered, the total variation in the table can be quantified by

$$X^2 = n \text{ trace} \left[\mathbf{Z} \mathbf{Z}' \right] = n \text{ trace} \left[\mathbf{Z}' \mathbf{Z} \right] \tag{2}$$

and has a chi-squared distribution with $(r - 1)(c - 1)$ degrees of freedom. Note that, for (1), $\mathbf{P} - \mathbf{rc}'$ is the difference between the observed proportions, $\mathbf{P}$, and its expected value under complete independence.

It is well known that the chi-squared statistic is proportional to the sample size. That is, even if the association structure between the row and column variables remains unchanged, the chi-squared statistic increases as the sample size increases. Therefore, correspondence analysis focuses on X^2/n as a measure of the strength of association between the variables and refers to it as the *total inertia* of the contingency table.

To determine the subspace needed to visually represent the association between the row and column variables, correspondence analysis is undertaken by applying SVD to the matrix of standardised residuals such that

$$D^* \left(\mathbf{D}_\lambda, \ \mathbf{U}, \ \mathbf{V} \right) = \mathbf{U} \mathbf{D}_\lambda \mathbf{V}'. \tag{3}$$

Here $\mathbf{U}$ is a $r \times M$ matrix and $\mathbf{V}$ is a $c \times M$ matrix where $M = \min(r, c) - 1$. These matrices contain column vectors which are row and column generalised basic vectors, respectively, and have the property

$$\mathbf{U}' \mathbf{U} = \mathbf{I}_M \qquad \mathbf{V}' \mathbf{V} = \mathbf{I}_M \tag{4}$$

where $\mathbf{I}_M$ is an $M \times M$ identity matrix. The matrix $\mathbf{D}_\lambda$ is a $M \times M$ diagonal matrix with elements $\lambda_1, \ldots, \lambda_M$ that are real and positive. They are the first M singular values of $\mathbf{Z}$ and are arranged in descending order so that

$$1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M > 0.$$

To obtain a graphical summary of the association using simple correspondence analysis, the variation between the row and column categories can be made by considering the matrix of coordinates

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{U} \mathbf{D}_\lambda \qquad\qquad \mathbf{G} = \mathbf{D}_c^{-1/2} \mathbf{V} \mathbf{D}_\lambda$$

such that

$$\frac{X^2}{n} = \text{trace} \left(\mathbf{F}' \mathbf{D}_r \mathbf{F} \right) = \text{trace} \left(\mathbf{G}' \mathbf{D}_c \mathbf{G} \right) = \text{trace} \left(\mathbf{D}_\lambda^2 \right).$$

Therefore row and column coordinates situated at some distance from the origin will coincide with large singular values, and hence a large total inertia. Similarly, a small total inertia will mean that the row and column coordinates will lie close to the centroid of the graphical display. This finding is consistent with all special cases considered in this paper.

Other types of graphical procedures may be considered, including the variety of different biplots that are available to the analyst. The following section provides

a generalised framework whereby this approach to correspondence analysis, non-symmetrical correspondence analysis and biplot analysis are special cases.

3 The unbalanced generalisation

3.1 The decomposition

The graphical representation of the association between the row and column responses in a low dimensional space is of interest here. Therefore we shall denote the $r \times M$ matrix of the row coordinates by $\mathbf{F}$ (as we did above). Similarly, denote the $c \times M$ matrix of column coordinates by $\mathbf{G}$. We can generalise the decomposition (1) by noting that D^* can be alternatively expressed in terms of these coordinates such that

$$\mathbf{Z} = D\left(\mathbf{D}_\lambda, \mathbf{F}, \mathbf{G}; \alpha, \beta\right) = \mathbf{W}_r^{-1}\mathbf{F}\mathbf{D}_\lambda^{1-(\alpha+\beta)}\mathbf{G}'\mathbf{W}_c^{-1} \tag{5}$$

for some value of α, and β where $\mathbf{F}$ and $\mathbf{G}$ may be expressed by

$$\mathbf{F} = \mathbf{W}_r\,\mathbf{U}\,\mathbf{D}_\lambda^\alpha, \tag{6}$$

$$\mathbf{G} = \mathbf{W}_c\,\mathbf{V}\,\mathbf{D}_\lambda^\beta \tag{7}$$

respectively. For (5), $\mathbf{W}_r$ is a $r \times r$ diagonal matrix of row weights and, depending on the analysis under consideration, may be specified as, for example, $\mathbf{D}_r^{-1/2}$ or by the identity matrix $\mathbf{I}_r$. Similarly $\mathbf{W}_c$ is a $c \times c$ diagonal matrix. The assignment of α and β can be chosen to be any value, or have any relationship, for example, $\alpha + \beta = 1$. Although α and β typically will be chosen so that $0 \le \alpha \le 1$ and $0 \le \beta \le 1$. As we shall see, typically, α (say) will be chosen to be either 0, 0.5 or 1 since the graphical display from such specifications will yield interpretable distances between row and column points.

By rearranging (6) and (7) and substituting them into (3) $\mathbf{Z}$ can be expressed in terms of $\mathbf{F}$, or $\mathbf{G}$, such that

$$\mathbf{Z} = \mathbf{W}_r^{-1}\mathbf{F}\mathbf{D}_\lambda^{1-\alpha}\mathbf{V}' = \mathbf{U}\mathbf{D}_\lambda^{1-\beta}\mathbf{G}'\mathbf{W}_c^{-1}. \tag{8}$$

Therefore, using (2) and (8), the total inertia is expressed in terms of $\mathbf{F}$ and $\mathbf{G}$ by

$$\frac{X^2}{n} = \text{trace}\left[\mathbf{D}_\lambda^{1-\alpha}\mathbf{F}'\mathbf{W}_r^{-2}\mathbf{F}\mathbf{D}_\lambda^{1-\alpha}\right] \tag{9}$$

$$= \text{trace}\left[\mathbf{D}_\lambda^{1-\beta}\mathbf{G}'\mathbf{W}_c^{-2}\mathbf{G}\mathbf{D}_\lambda^{1-\beta}\right]. \tag{10}$$

In fact, by substituting $\mathbf{F}$ and $\mathbf{G}$, as defined by (6) and (7), back into (9) and (10), we can see that

$$\frac{X^2}{n} = \text{trace}\left(\mathbf{D}_\lambda^2\right).$$

That is, the total inertia of the contingency table is measured as the sum of squares of the singular values obtained from the SVD of the matrix of standardised residuals, $\mathbf{Z}$. This result also shows that, irrespective of the choice of α, β, $\mathbf{W}_r$ and $\mathbf{W}_c$, the mth principal axis of the graphical display will account for $100 \times \left(\lambda_m^2 / \frac{X^2}{n} \right)$ % of the variation that exists in the contingency table. So from a descriptive point of view the interpretation of the axes will be identical, but the relative position, and configuration, of the row and column coordinates will depend on the choice of α, β, $\mathbf{W}_r$ and $\mathbf{W}_c$.

Consider the matrix of row coordinates, $\mathbf{F}$, given by (6). We can express $\mathbf{U}$ in terms of this matrix such that

$$\mathbf{U} = \mathbf{W}_r^{-1} \mathbf{F} \mathbf{D}_\lambda^{-\alpha}$$

so that, from its orthogonality property, we obtain the result

$$\mathbf{D}_\lambda^{2\alpha} = \mathbf{F}' \mathbf{W}_r^{-2} \mathbf{F}. \tag{11}$$

Therefore, since the right hand side of (11) is diagonal, the total inertia may be expressed in terms of $\mathbf{F}$ by

$$\frac{X^2}{n} = \text{trace} \left[\left(\mathbf{F}' \mathbf{W}_r^{-2} \mathbf{F} \right)^{1/\alpha} \right] \tag{12}$$

for $\alpha = 1$. When the total inertia of the contingency table remains unchanged, and keeping $\mathbf{W}_r$ constant, if α increases then (12) shows that the configuration of row coordinates in the low-dimensional space will be situated some distance from the centre of the display, called the centroid. If, on the other hand, α is relatively small (close to zero, say) then the row coordinates will be drawn towards the centroid. Similarly, when $\beta = 1$, by using the orthogonality property of $\mathbf{V}$, the total inertia can be expressed in terms of the column coordinates $\mathbf{G}$ by

$$\frac{X^2}{n} = \text{trace} \left[\left(\mathbf{G}' \mathbf{W}_c^{-2} \mathbf{G} \right)^{1/\beta} \right]$$

and so the configuration of column coordinates relative to the centroid will depend on the choice of β.

3.2 Reconstituting P

An important feature of the generalisation (5) is that the matrix of joint proportions, $\mathbf{P}$ can be reconstituted from the row and column coordinates, the generalised basic vectors and the singular values.

If the matrices $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{D}_\lambda$ are of full rank, then the values of $\mathbf{P}$ can be reconstituted exactly by

$$\mathbf{P} = \mathbf{r}\mathbf{c}' + \mathbf{D}_r^{1/2} \mathbf{Z} \mathbf{D}_c^{1/2}.$$

This is termed the *saturated* reconstitution formula and can be alternatively expressed in terms of the matrices of row and column coordinates, $\mathbf{F}$ and $\mathbf{G}$, by

$$\mathbf{P} = \mathbf{r}\mathbf{c}' + \tilde{\mathbf{W}}_r \mathbf{F} \mathbf{D}_\lambda^{1-(\alpha+\beta)} \mathbf{G}' \tilde{\mathbf{W}}_c \tag{13}$$

where $\tilde{\mathbf{W}}_r = \mathbf{D}_r^{1/2} \mathbf{W}_r^{-1}$ and $\tilde{\mathbf{W}}_c = \mathbf{W}_c^{-1} \mathbf{D}_c^{1/2}$. Therefore $\tilde{\mathbf{W}}_r \mathbf{F} \mathbf{D}_\lambda^{1-(\alpha+\beta)} \mathbf{G}' \tilde{\mathbf{W}}_c$ measures the deviation from complete independence of the cell proportions. Complete independence between the two categorical variables will result when all $\lambda_m = 0$ regardless of the choice of α or β. When this is the case, by (11), the row coordinates will be situated at the centroid. Similarly, the column coordinates will also exist at the centroid.

The advantage of the reconstitution formula is that, even if a subset of axes are selected to graphically depict the association, the joint proportions in $\mathbf{P}$ can be approximately reconstituted. This is helpful when a test of the number of dimensions needed to construct the plot is required. For example, Escoufier (1988) looks at this issue using a special case of (13) for correspondence analysis; see Sect. 6.3

Suppose $\mathbf{F}_{(k)}$ and $\mathbf{G}_{(k)}$ contain the first k column vectors of $\mathbf{F}$ and $\mathbf{G}$, respectively, so that they are the row and column coordinates for the first $k < M$ dimensions of a joint plot. Also suppose that $\mathbf{D}_{\lambda(k)}$ is the $k \times k$ diagonal matrix of the first k singular values. Then $\mathbf{P}$ can be approximated by the *unsaturated* reconstitution model

$$\mathbf{P} \approx \mathbf{P}_{(k)} = \mathbf{r}\mathbf{c}' + \tilde{\mathbf{W}}_r \mathbf{F}_{(k)} \mathbf{D}_{\lambda(k)}^{1-(\alpha+\beta)} \mathbf{G}'_{(k)} \tilde{\mathbf{W}}_c . \tag{14}$$

3.3 Transition formula

Another special property of the general decomposition (5) is that one can obtain the row coordinates from the column coordinates, denoted by $\mathbf{G} \to \mathbf{F}$, and vice versa, that is $\mathbf{F} \to \mathbf{G}$. Formulae which describe such transitions are called *transition formulae*.

By rearranging (6) and (7), the transition formula for $\mathbf{G} \to \mathbf{F}$ is

$$\mathbf{F} = \mathbf{W}_r \mathbf{Z} \mathbf{W}_c^{-1} \mathbf{G} \mathbf{D}_\lambda^{\alpha-\beta-1}$$

while the formula for $\mathbf{F} \to \mathbf{G}$ is

$$\mathbf{G} = \mathbf{W}_c \mathbf{Z}' \mathbf{W}_r^{-1} \mathbf{F} \mathbf{D}_\lambda^{\beta-\alpha-1} .$$

As a result of these formulae, cells with a large standardised residual will result in the row and column responses that cross-classify that cell to be at close proximity to one another.

Let's consider some popular coordinate systems that are specific cases of these generalisations.

4 Special unbalanced cases

4.1 $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}_c$ and $\alpha + \beta = 1$

With this specification of $\mathbf{W}_r$, $\mathbf{W}_c$, α and β we obtain the matrix of residuals

$$\mathbf{Z} = \mathbf{F}\,\mathbf{G}'$$

which is the factorisation, for categorical data, used for the construction of the biplot; see Gabriel (1971). Thus, the row and column coordinates are

$$\mathbf{F} = \mathbf{U}\mathbf{D}_\lambda^\alpha$$
$$\mathbf{G} = \mathbf{V}\mathbf{D}_\lambda^{1-\alpha}.$$

Coordinates of this form were described by, for example, Aitchison and Greenacre (2002) and Goodman (1996, pp. 422–424). Different values of α will produce different coordinates. The choice of α, and therefore β, is made depending on the variable of interest. This is an important issue when meaningful interpretations are required and depend on the variable of interest. For example, for the analysis of quantitative variables (Gabriel 1971, p. 458) describes three sets coordinates which coincide with $\alpha = 1/2$, $\alpha = 1$ and $\alpha = 0$ so that $\beta = 1/2$, $\beta = 0$ and $\beta = 1$ respectively. In the case of categorical variables, the assignment of these values have also been described as *symmetric*, *row isometric* and *column isometric* factorisations, respectively; refer to Lombardo et al. (1996). The case where $\alpha = 0$ and $\beta = 1$ was discussed by Grassi and Visentin (1994) who refer to their graphical displays using these sets of coordinates as "a symmetric β-plots".

Special characteristics of this analysis include expressing the total inertia by

$$\frac{X^2}{n} = \text{trace}\left[\mathbf{D}_\lambda^{1-\alpha}\mathbf{F}'\mathbf{F}\mathbf{D}_\lambda^{1-\alpha}\right] = \text{trace}\left[\mathbf{D}_\lambda^\alpha\mathbf{G}'\mathbf{G}\mathbf{D}_\lambda^\alpha\right].$$

Also, the matrix of joint probabilities $\mathbf{P}$ can be reconstituted by

$$\mathbf{P} = \mathbf{rc}' + \mathbf{D}_r^{-1/2}\mathbf{F}\,\mathbf{G}'\mathbf{D}_c^{-1/2} \tag{15}$$

while the transition formulae are

$$\mathbf{F} = \mathbf{D}_r^{-1/2}\mathbf{P}\mathbf{D}_c^{-1/2}\mathbf{G}\mathbf{D}_\lambda^{2(\alpha-1)}$$
$$\mathbf{G} = \mathbf{D}_c^{-1/2}\mathbf{P}'\mathbf{D}_r^{-1/2}\mathbf{F}\mathbf{D}_\lambda^{-2\alpha}.$$

Thus, for the row isometric factorisation $\mathbf{F} = \mathbf{D}_r^{-1/2}\mathbf{P}\mathbf{D}_c^{-1/2}\mathbf{G}$ and, for the column isometric factorisation, $\mathbf{G} = \mathbf{D}_r^{-1/2}\mathbf{P}'\mathbf{D}_r^{-1/2}\mathbf{F}$.

If we considered instead the specification $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$ and $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$, then the row and column principal coordinates are defined as

$$\mathbf{F} = \mathbf{D}_r^{-1/2}\mathbf{U}\mathbf{D}_\lambda^\alpha$$
$$\mathbf{G} = \mathbf{D}_c^{-1/2}\mathbf{V}\mathbf{D}_\lambda^{1-\alpha}$$

and have been the subject of much discussion in the correspondence analysis literature. These coordinates are described by Lorenzo-Seva et al. (2009) and Velden and Kiers (2005) where their focus was on the rotation of the configuration of points to yield the most interpretable display. Just as others have done in similar situations, Lorenzo-Seva et al. (2009) and Velden and Kiers (2005) described the advantages of considering specific values of α; those being $\alpha = 0$, $\alpha = 1$ and $\alpha = 1/2$. Such a generalisation (irrespective of the choice of α) leads to the same type of biplots described above and so the reader is invited to consider these articles for a more indepth discussion of the issue. For a general discussion of the role of biplots in correspondence analysis one may refer to Greenacre (1993), Gower et al. (2011, 2014). See also Section 4.5.3 of Beh and Lombardo (2014) for an overview of a variety of biplot displays for correspondence analysis.

4.2 $\mathbf{W_r} = n^{-1/2}\mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{1/2}$, $\alpha = 1$ and $\beta = 0$

For this combination of values, the row and column coordinates are

$$\mathbf{F} = n^{-1/2}\mathbf{D}_r^{-1/2}\,\mathbf{U}\,\mathbf{D}_\lambda \tag{16}$$
$$\mathbf{G} = \mathbf{D}_c^{1/2}\mathbf{V} \tag{17}$$

which is the coordinate system used for the column metric preserving (CMP) biplot for categorical variables; see Gabriel and Orodoff (1990, p. 483). In this case, the matrix of standardised residuals may be expressed in terms of $\mathbf{F}$ and $\mathbf{G}$ by

$$\mathbf{Z} = n^{-1/2}\mathbf{D}_r^{1/2}\mathbf{F}\mathbf{G}'\mathbf{D}_c^{-1/2}$$

which leads to the singular value decomposition of the centred row profiles $\mathbf{D}_r^{-1}\mathbf{P} - \mathbf{1c}'$. For a CMP analysis, the saturated reconstitution formula is

$$\mathbf{P} = \mathbf{rc}' + n^{-1/2}\mathbf{D}_r\mathbf{F}\mathbf{G}'$$

and the transition formulae $\mathbf{G} \rightarrow \mathbf{F}$ and $\mathbf{F} \rightarrow \mathbf{G}$ are

$$\mathbf{F} = n^{-1/2}\mathbf{D}_r^{-1}\mathbf{P}\mathbf{D}_c^{-1}\mathbf{G}$$
$$\mathbf{G} = n^{1/2}\mathbf{P}'\mathbf{F}\mathbf{D}_\lambda^{-2}$$

respectively. Note that, for $\mathbf{F} \rightarrow \mathbf{G}$, $\mathbf{D}_r^{-1}\mathbf{P}\mathbf{D}_c^{-1}$ is just the matrix of Pearson ratio's; see Beh (2004), Beh and Lombardo (2014) and Goodman (1996). For such ratio's,

complete independence between the row and column variables will arise when all the elements of $\mathbf{D}_r^{-1}\mathbf{P}\mathbf{D}_c^{-1}$ are equal to 1. Therefore in this case we get, under complete independence, $\mathbf{P} = \mathbf{rc}'$, as expected.

5 A balanced generalisation

A generalisation that leads to other popular approaches is when, for (5), $\alpha = \beta$. Since, for our general framework, only the matrix of row and column coordinates, $\mathbf{F}$ and $\mathbf{G}$, involve α and β, such a result means that the association that exists between the variables forming a contingency table is being represented equally by the row and column categories. Therefore one may consider such a framework to be balanced. The nature of the representation of the row and column coordinates may be considered through the specification of $\mathbf{W}_r$ and $\mathbf{W}_c$. A balanced analysis may further be considered by specifying these matrices such that $\mathbf{W}_r = \mathbf{D}_r$ and $\mathbf{W}_c = \mathbf{D}_c$, or $\mathbf{W}_r = \mathbf{I}_r$ and $\mathbf{W}_c = \mathbf{I}_c$, say.

By imposing the constraint $\alpha = \beta$, the matrix of standardised residuals, $\mathbf{Z}$, can be expressed by

$$\mathbf{Z} = D\left(\mathbf{D}_\lambda, \mathbf{F}, \mathbf{G}; \alpha\right) = \mathbf{W}_r^{-1}\mathbf{F}\mathbf{D}_\lambda^{1-2\alpha}\mathbf{G}'\mathbf{W}_c^{-1} \tag{18}$$

for $0 \le \alpha \le 1$, so that the row and column coordinates are

$$\mathbf{F} = \mathbf{W}_r\,\mathbf{U}\,\mathbf{D}_\lambda^\alpha \tag{19}$$
$$\mathbf{G} = \mathbf{W}_c\,\mathbf{V}\,\mathbf{D}_\lambda^\alpha \tag{20}$$

respectively. Using these results, the total inertia can be expressed by

$$\begin{aligned}\frac{X^2}{n} &= \text{trace}\left[\mathbf{D}_\lambda^{1-\alpha}\mathbf{F}'\mathbf{W}_r^{-2}\mathbf{F}\mathbf{D}_\lambda^{1-\alpha}\right]\\ &= \text{trace}\left[\mathbf{D}_\lambda^{1-\alpha}\mathbf{G}'\mathbf{W}_c^{-2}\mathbf{G}\mathbf{D}_\lambda^{1-\alpha}\right]\\ &= \text{trace}\left(\mathbf{D}_\lambda^2\right).\end{aligned} \tag{21}$$

For the balanced generalisation the reconstitution of $\mathbf{P}$ can be made so that

$$\mathbf{P} = \mathbf{rc}' + \tilde{\mathbf{W}}_r\mathbf{F}\mathbf{D}_\lambda^{1-2\alpha}\mathbf{G}'\tilde{\mathbf{W}}_c$$

and the transition formulae for $\mathbf{G} \to \mathbf{F}$ and $\mathbf{F} \to \mathbf{G}$ are

$$\mathbf{F} = \mathbf{W}_r\mathbf{Z}\mathbf{W}_c^{-1}\mathbf{G}\mathbf{D}_\lambda^{-1}$$
$$\mathbf{G} = \mathbf{W}_c\mathbf{Z}'\mathbf{W}_r^{-1}\mathbf{F}\mathbf{D}_\lambda^{-1}$$

respectively. Thus, for the balanced generalisation, the transition formulae do not depend on the value of α, but only rely on the choice of $\mathbf{W}_r$ and $\mathbf{W}_c$.

6 Special balanced cases

6.1 $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$, $\alpha = 0$

One of the most simple coordinate system that we can use is when $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 0$. The matrix of standardised residuals is

$$\mathbf{Z} = \mathbf{D}_r^{1/2} \mathbf{F} \mathbf{D}_\lambda \mathbf{G}' \mathbf{D}_c^{1/2}$$

where the row and column coordinates are

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{U}$$
$$\mathbf{G} = \mathbf{D}_c^{-1/2} \mathbf{V}$$

respectively. These coordinates are commonly referred to as standard coordinates since each of the axes that are used to construct the plot have an equal weight of 1; see, for example, Beh and Lombardo (2014) and Greenacre (1984). This is evident since, by setting $\alpha = 0$ in (11),

$$\mathbf{D}_\lambda^{2\alpha} = \mathbf{F}' \mathbf{D}_r \mathbf{F} = \mathbf{U}' \mathbf{U} = \mathbf{I}_M.$$

In this case the saturated reconstitution formulae is

$$\mathbf{P} = \mathbf{rc}' + \mathbf{D}_r \mathbf{F} \mathbf{D}_\lambda \mathbf{G}' \mathbf{D}_c$$
$$= \mathbf{rc}' \left(\mathbf{1} + \mathbf{F} \mathbf{D}_\lambda \mathbf{G}' \right)$$

where here $\mathbf{1}$ is an $r \times c$ matrix of 1's.

6.2 $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 1$

For this combination of $\mathbf{W}_r$, $\mathbf{W}_c$ and α, the matrix of residuals may be decomposed so that

$$\mathbf{Z} = \mathbf{D}_r^{1/2} \mathbf{F} \mathbf{D}_\lambda^{-1} \mathbf{G}' \mathbf{D}_c^{1/2}$$

where the row and column coordinates can be expressed by

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{U} \mathbf{D}_\lambda$$
$$\mathbf{G} = \mathbf{D}_c^{-1/2} \mathbf{V} \mathbf{D}_\lambda$$

respectively. These coordinates are those obtained from the classical version of the simple correspondence analysis of a two-way contingency table; see, for example,

Beh (2004), Beh and Lombardo (2014), Greenacre (1984) and Greenacre (2010b). Therefore, the total inertia can be expressed in terms of $\mathbf{F}$ and $\mathbf{G}$ such that

$$\frac{X^2}{n} = \text{trace}\left[\mathbf{F}'\mathbf{D}_r\mathbf{F}\right] = \text{trace}\left[\mathbf{G}'\mathbf{D}_c\mathbf{G}\right]$$

which is the trace of the simple weighted Euclidean distances between the row coordinates (and column coordinates). The reconstitution formula is, as it is for simple correspondence analysis,

$$\mathbf{P} = \mathbf{rc}' + \mathbf{D}_r\mathbf{F}\mathbf{D}_\lambda^{-1}\mathbf{G}'\mathbf{D}_c \tag{22}$$
$$= \mathbf{rc}' \left(1 + \mathbf{F}\mathbf{D}_\lambda^{-1}\mathbf{G}'\right) \tag{23}$$

and is also referred to as correspondence model by Goodman (1986) and Greenacre (1984). It was pointed out by Gilula et al. (1988) that there are important differences between the interpretation of the elements of $\mathbf{D}_\lambda$ for the correspondence model and the Goodman RC model. Refer to these authors for more details. This highlights that careful interpretation of the configuration of points is needed when looking at different choices of weights in (5) and (18).

Using this combination of α, $\mathbf{W}_r$ and $\mathbf{W}_c$ leads to the transition formulae

$$\mathbf{F} = \mathbf{D}_r^{-1}\mathbf{P}\mathbf{G}\mathbf{D}_\lambda^{-1}$$
$$\mathbf{G} = \mathbf{D}_c^{-1}\mathbf{P}'\mathbf{F}\mathbf{D}_\lambda^{-1}.$$

Note that this approach to performing simple correspondence analysis is equivalent to considering Pearson's contingencies $\mathbf{D}_r^{-1}\mathbf{P}\mathbf{D}_c^{-1} - \mathbf{1}$; see Goodman (1996). However, in this case $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 1/2$.

6.3 $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}_c$, $\alpha = 1/2$

When $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}_c$, $\alpha = 1/2$, the decomposition of the matrix of standardised residual takes the form

$$\mathbf{Z} = \mathbf{F}\mathbf{G}'$$

and is a special case of the biplot generalisation of (5) where $\alpha = \beta = 1/2$ so that $\alpha + \beta = 1$. Therefore row and column coordinates

$$\mathbf{F} = \mathbf{U}\mathbf{D}_\lambda^{1/2}$$
$$\mathbf{G} = \mathbf{V}\mathbf{D}_\lambda^{1/2}$$

are *balanced* factorisations of the matrix of residual, while the saturated reconstitution formulae is given by (15). These coordinates can be confirmed by substituting $\alpha = 1/2$

into (19) and (20). For this case the total inertia may be expressed in terms of the row and column coordinates by

$$\frac{X^2}{n} = \text{trace}\left[\mathbf{D}_\lambda \mathbf{G}'\mathbf{G}\right] = \text{trace}\left[\mathbf{D}_\lambda \mathbf{F}'\mathbf{F}\right]$$

Also the matrix of singular values can be expressed as

$$\mathbf{D}_\lambda = \mathbf{F}'\mathbf{F} = \mathbf{G}'\mathbf{G}$$

so that (21) is verified. For this specification of $\mathbf{W}_r$, $\mathbf{W}_c$ and α, the transition formulae are

$$\mathbf{F} = \mathbf{D}_r^{-1/2}\mathbf{P}\mathbf{G}$$
$$\mathbf{G} = \mathbf{D}_r^{-1/2}\mathbf{P}'\mathbf{F}.$$

Therefore, a relatively large cell value will mean that the row and column categories that share this cell will be drawn close together in the display.

7 Simple demonstration

7.1 The data

We now consider a simple demonstration of the generalisation using the data summarised in the two-way contingency table given by Table 1. These data were used as a motivating example in the exploration of correspondence analysis by Beh and Lombardo (2014). It cross-classifies 1117 insulation workers in metropolitan New York by the number of years they have been exposed to asbestos and the severity of asbestosis they were diagnosed with. This contingency table appears in Beh and Smith (2011) but is from a more comprehensive study of the the impact of asbestos on health by Selikoff (1981). We note that Beh and Lombardo (2014) did not examine the association between the two variables of Table 1 in the context of any generalised form of correspondence analysis. Therefore, we compare the results of the analysis of Table 1 under the following two special cases:

– *Unbalanced case* $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}_c$, $\alpha = 0.9$ and $\beta = 0.1$, which coincides with the special unbalanced case outlined in Sect. 4.1 (for a specific value of α and β where $\alpha + \beta = 1$)
– *Balanced case* $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 1$, which coincides with the special balanced case outlined in Sect. 4.2.

Any of the other special cases could also have been considered but we restrict our attention to these two cases for the sake of brevity. The two analyses performed here is made using the R function `gensimpleca()` given in the electronic supplementary material that accompanies this paper. Before we consider the application of the generalisation outlined above, we briefly examine the association between the variables. For Table 1, the chi-squared statistic, X^2 is 648.812 and has a p-value that is

Table 1 Selikoff's data for 1117 New York workers with occupational exposure to asbestos

Occupational exposure (years)	Asbestos grade diagnosed				Total
	None	Grade 1	Grade 2	Grade 3	
0–9	310	36	0	0	346
10–19	212	158	9	0	379
20–29	21	35	17	4	77
30–39	25	102	49	18	194
40+	7	35	51	28	121
Total	575	366	126	50	1117

less than 0.0001. Therefore, there is a statistically significant association between the number of years of occupational exposure to asbestos and the severity of asbestosis that a worker contracts. From a correspondence analysis perspective, the total inertia is $648.812/1117 = 0.581$. The nature of the association can be explored using the traditional approach to simple correspondence analysis. Although we shall instead consider here the two special cases of the generalisation presented above, where the second case is equivalent to simple correspondence analysis. To visualise the association between the row and column categories of Table 1, the optimal number of dimensions from the generalisation is $\min(5, 4) - 1 = 3$ and so the total inertia can be partitioned to yield a contribution that each axis makes. We shall show that, from a practical perspective using the two cases described above, both yield identical numerical summaries of the total inertia and principal inertia values and many other features.

7.2 *Unbalanced case*: $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}_c$, $\alpha = 0.9$ and $\beta = 0.1$

For this specification of parameters, the two-dimensional display is given by Fig. 1. This plot suggests that workers who were exposed to asbestos for no more than 10 years are likely to not be diagnosed with any form of asbestosis, while those with more than 40 years of exposure are linked to the highest levels of severity.

From a numerical perspective, the overall association can be quantified by the total inertia of the contingency table using (12). Doing so gives $X^2/n = 0.581$ which is equivalent to the total inertia calculated in Sect. 7.1; see also the numerical output summarised in the Appendix.

For each axis of the optimal three-dimensional plot, Table 2 provides a summary of the principal inertia values, their percentage contribution to the total inertia and the cummulative inertia. It shows that a three-dimensional plot, defined by specifying the above combination of $\mathbf{W}_r$, $\mathbf{W}_c$, α and β, visually represents 100% of the association between the variables- this should be of no surprise. It also shows that the two-dimensional plot of Fig. 1 accounts for 99.567% of this association - which makes it a very good quality visual summary of the association. We can also reconstitute the cell frequencies of the contingency table using the reconstitution formula (14), where $k = 2$. Doing so gives expected cell frequencies summarised in Table 3.
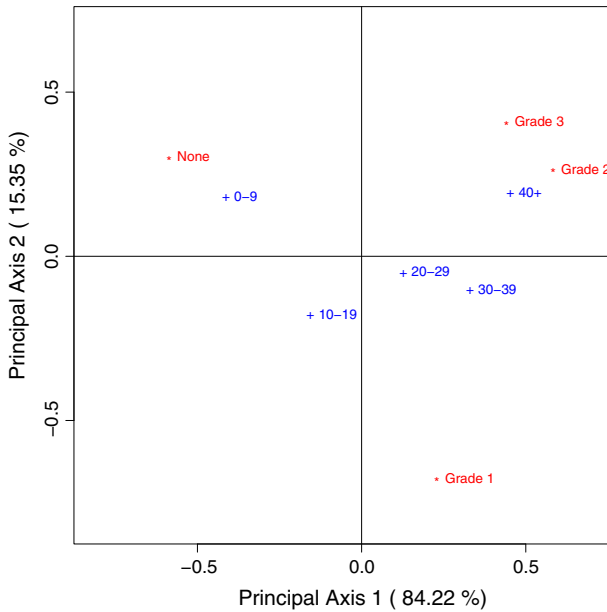
Fig. 1 A two-dimensional display of the association between the variables of Table 1 where $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}$, $\alpha = 0.9$ and $\beta = 0.1$

Table 2 Summary of the principal inertia, its percentage contribution to the total inertia and the cummulative percentage contribution to the total inertia

Axis	Inertia	% Inertia	Cumm. % inertia
1	0.489	84.215	84.215
2	0.089	15.352	99.567
3	0.003	0.433	100.000

Table 3 Two-dimensional reconstitution of Table 1

Occupational exposure (years)	Asbestos grade diagnosed				Total
	None	Grade 1	Grade 2	Grade 3	
0–9	309.807	36.644	− 1.498	1.047	346.000
10–19	212.471	158.433	12.645	− 2.549	379.000
20–29	20.699	36.003	14.666	5.632	77.000
30–39	24.837	102.541	47.741	18.880	193.999
40+	7.187	35.379	52.445	26.990	121.001
Total	575.001	366.000	125.999	50.000	1117.000

7.3 Balanced case: $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 1$

Suppose we now consider the specification of parameters described in Sect. 4.2. That is, substitute $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 1$ into (18)–(20). In doing so we
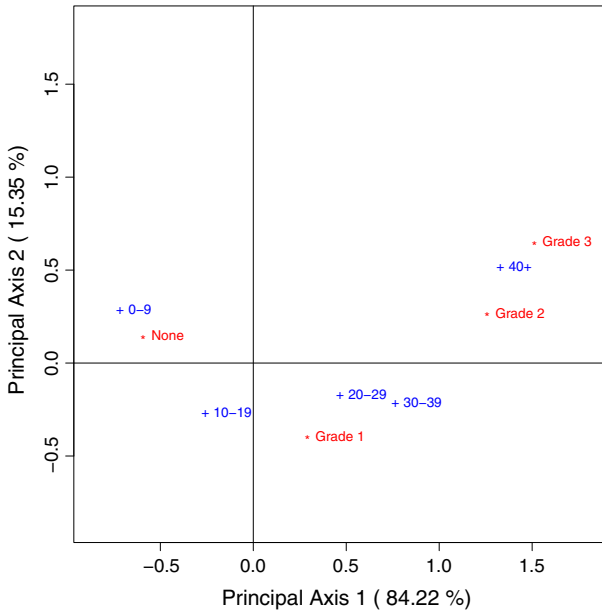
Fig. 2 A two-dimensional display of the association between the variables of Table 1 where $\mathbf{W}_r = \mathbf{D}_r^{-1/2}$, $\mathbf{W}_c = \mathbf{D}_c^{-1/2}$ and $\alpha = 1$

obtain the same total inertia value and principal inertia values calculated in Sect. 7.2. Therefore, the two special cases yield identical percentage contributions to the total inertia for each of the three axes; see Table 2 for these summaries. While these numerical summaries are equivalent to those given in Sect. 7.2, this specification of parameters yields the same two-dimensional reconstituted cell frequencies that appear in Table 3. However, since the scaling of the row and column points are different (due to the different specification of $\mathbf{W}_r$, $\mathbf{W}_c$ and α) the two-dimensional visual display provides a slightly different configuration of points when compared with Fig. 1; Fig. 2 provides a two-dimensional display that is just the traditional two-dimensional correspondence plot obtained from performing simple correspondence analysis on Table 1.

8 Other issues

8.1 Non-symmetrical correspondence analysis

Further generalisations can be made by considering other residuals or decompositions of $\mathbf{Z}$. For example, they can be modified to include as a special case non-symmetrical correspondence analysis (NSCA); see, for example, D'Ambra and Lauro (1989, 1992) and Kroonenberg and Lombardo (1998). In the case where the column categories are treated as responses to an explanatory variable and the row categories comprising of the response variable, the matrix of standardised residuals, (1), can be modified to consider NSCA by considering

$$\tilde{\mathbf{Z}} = \left(\mathbf{P} - \mathbf{rc}'\right)\mathbf{D}_c^{-1/2} = \mathbf{U}\mathbf{D}_\lambda\mathbf{V}'.$$

As a result, $N_\tau = n\,\text{trace}\left[\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}}\right]$ is the numerator of the Goodman-Kruskal tau index, Goodman and Kruskal (1954), and represents the absolute increase in predictability of a row response given the knowledge of a column category. Therefore, NSCA can be viewed as an example of the column isometric factorisation of $\tilde{\mathbf{Z}}$ (where $\mathbf{W}_r = \mathbf{I}_r$, $\mathbf{W}_c = \mathbf{I}_c$, $\alpha = 0$ and $\beta = 1$) so that $\mathbf{F} = \mathbf{U}$ and $\mathbf{G} = \mathbf{V}\mathbf{D}_\lambda$ and $N_\tau = \text{trace}\left[\mathbf{D}_\lambda^2\right] = \text{trace}\left[\mathbf{G}'\mathbf{G}\right]$; see also Sect. 4.1. This suggests that a further generalisation may be considered by replacing $\mathbf{Z}$ with $\bar{\mathbf{Z}}$ in (1) where

$$\bar{\mathbf{Z}} = \mathbf{D}_r^{-\phi/2}\left(\mathbf{P} - \mathbf{rc}'\right)\mathbf{D}_c^{-\psi/2}. \tag{24}$$

Therefore, if we let $\hat{\mathbf{Z}} = \hat{\mathbf{Z}}_{(\phi,\,\psi)}$ the general total inertia of the contingency table may be expressed as $\text{trace}\left(\hat{\mathbf{Z}}'_{(\phi,\,\psi)}\hat{\mathbf{Z}}_{(\phi,\,\psi)}\right)$. The value of the inertia depends on the analysis conducted and therefore depends on the specification of ϕ and ψ. For example, the following specifications of ϕ and ψ lead to the following types of analyses

- $\phi = 1$ and $\psi = 1$ - the correspondence analysis, or biplot analysis, of the matrix of standardised residuals,
- $\phi = 0$ and $\psi = 1$ - the NSCA of a contigency table where the rows are response categories and the columns are explanatory categories,
- $\phi = 1$ and $\psi = 0$ - the NSCA of a contingency table where the rows are explanatory categories and the columns are response categories.

The way in which the graphical depiction of association is represented may be obtained using those specifications of $\mathbf{W}_r$, $\mathbf{W}_c$, α and β given above.

8.2 Analysis of adjusted residuals

As we have discussed throughout this paper, typically, the graphical display of association using methods akin to correspondence analysis and biplots, involves the SVD of Pearson's standardised residuals; see (1). However, Agresti (2002, p. 81) points out that the asymptotic variance of the residuals are less than 1, and Haberman (1973) notes that the variance of the (i, j)th cell of $n^{-1/2}\mathbf{Z}$ is $(1 - p_{i\bullet})(1 - p_{\bullet j})$ where $p_{i\bullet}$ and $p_{\bullet j}$ are the (i, i)th and (j, j)th element of $\mathbf{D}_r$ and $\mathbf{D}_c$ respectively. Therefore, one may consider the matrix of adjusted residuals

$$\hat{\mathbf{Z}} = \mathbf{D}_r^{-1/2}\left(\mathbf{I}_r - \mathbf{D}_r\right)^{-1/2}\left(\mathbf{P} - \mathbf{rc}'\right)\left(\mathbf{I}_c - \mathbf{D}_c\right)^{-1/2}\mathbf{D}_c^{-1/2}$$

which can also be incorporated into the generalisation. Recently, Beh (2012) considered the application of correspondence using adjusted residuals. Therefore, such an adaptation of correspondence analysis may be incorporated into the general framework discussed here by considering the SVD of the adjusted residuals a special case of

$$\hat{\mathbf{Z}} = \mathbf{D}_r^{-\phi/2} \left(\mathbf{I}_r - \mathbf{D}_r\right)^{-\gamma/2} \left(\mathbf{P} - \mathbf{rc}'\right) \left(\mathbf{I}_c - \mathbf{D}_c\right)^{-\delta/2} \mathbf{D}_c^{-\psi/2} \tag{25}$$

where $\phi = \psi = \gamma = \delta = 1$. Decompositions akin to (5) and (18) may therefore be adopted by considering $\hat{\mathbf{Z}}$ instead of $\mathbf{Z}$. One may note that if $\gamma = \delta = 0$, then (25) is just the matrix of residuals given by (24).

8.3 Freeman–Tukey residual

Yet another generalisation of our approach is that, rather than considering the standardised residual, or its adjusted version, correspondence analysis can be performed through the SVD of the Freeman–Tukey residual. The matrix of Freeman–Tukey residuals is of the form

$$\check{\mathbf{Z}} = 2 \left(\mathbf{P}^{1/2} - \left(\mathbf{rc}'\right)^{1/2}\right)$$

so that the total inertia using the Freeman–Tukey statistic (Freeman and Tukey 1950) is

$$\frac{T^2}{n} = \mathrm{trace}\left[\check{\mathbf{Z}}'\check{\mathbf{Z}}\right] = \mathrm{trace}\left[\check{\mathbf{Z}}\check{\mathbf{Z}}'\right].$$

The role of this statistic in correspondence analysis was briefly described by Beh and Lombardo (2014, Section 9.3). So, the residual $\check{\mathbf{Z}}$ can be built into the general residual of (25) by

$$\hat{\mathbf{Z}} = 2^{\theta-1} \left[\mathbf{D}_r^{-\phi/2} \left(\mathbf{I}_r - \mathbf{D}_r\right)^{-\gamma/2}\right]^{\theta-1} \left(\mathbf{P}^{\theta/2} - \left(\mathbf{rc}'\right)^{\theta/2}\right)$$
$$\times \left[\left(\mathbf{I}_c - \mathbf{D}_c\right)^{-\delta/2} \mathbf{D}_c^{-\psi/2}\right]^{\theta-1}.$$

When $\theta = 2$ this expression reduces to (25) while the matrix of Freeman–Tukey residuals when $\theta = 1$.

8.4 Multiple correspondence analysis

The algebraic generalisations given in this paper have been described with the analysis of two categorical variables in mind. However, the same generalisations can also be applied to the analysis of multiple categorical variables in certain cases. These cases are when the traditional approach to multiple correspondence analysis is performed and this arises when the multi-way contingency table is converted to either its indicator matrix form or Burt matrix form, then simple correspondence analysis is applied to these matrices. For more information on multiple correspondence analysis, see, for example, Beh and Lombardo (2014), Greenacre (1984, 1990) and Greenacre and Blasius (2006).

9 Discussion

This paper presents a simple algebraic generalisation that unifies the growing number of variations and features of the correspondence analysis of a two-way contingency table. These approaches are reflected by the unbalanced equation (5) and its balanced version (18). By considering (25), the association structure of the row and column categories may be investigated by considering the departure from complete independence via the more general result

$$\mathbf{P} = \mathbf{rc}' + \bar{\mathbf{W}}_r \mathbf{FD}_\lambda^{1-(\alpha+\beta)} \mathbf{G}' \bar{\mathbf{W}}_c'$$

where

$$\bar{\mathbf{W}}_r = \mathbf{D}_r^{\phi/2} \left(\mathbf{I}_r - \mathbf{D}_r\right)^{\gamma/2} \mathbf{W}_r^{-1}$$
$$\bar{\mathbf{W}}_c = \mathbf{D}_r^{\psi/2} \left(\mathbf{I}_c - \mathbf{D}_c\right)^{\delta/2} \mathbf{W}_c^{-1} .$$

For such a specification, the row and column categories may be graphically represented by considering their matrix of coordinates

$$\mathbf{F} = \mathbf{W}_r \mathbf{UD}_\lambda^\alpha ,$$
$$\mathbf{G} = \mathbf{W}_c \mathbf{VD}_\lambda^\beta$$

respectively.

A key issue of this framework then is to decide on the specification of $\alpha, \beta, \gamma, \delta, \phi, \psi, \mathbf{W}_r$ and $\mathbf{W}_c$. By considering specific quantifications of these, one can undertake a variety of known analyses that fall within the correspondence analysis (and related biplot) family of techniques. One may consider, for example, any one of the special cases described above. The framework described in this paper also opens up the possibility of considering previously unexplored options such as implementing NSCA using an adaptation of adjusted residuals (although the benefits of doing so will not be investigated here).

Our generalisation, and much of the literature that discusses the construction, choice and interpretation of biplots, focuses on specific values of α and β; namely, for $\alpha, \alpha = 0$ and $\alpha = 1$. These values of α lead to interpretations of these asymmetric biplots that are relatively simple. This is because the categories of one variable are depicted using the traditional principal coordinates while the other variable is depicted using standard coordinates. Such specifications of α allow for row/column distances to be meaningful and, through the principal coordinates, for the reconstruction of the total inertia of the contingency table. However, there is no reason why alternative specifications of α (and β) can not be made, although in such cases the total inertia may well be an unknown measure of association. We, therefore, leave further investigations of this topic for future consideration.

The generalisation described above are not confined to the implementation of SVD. Other methods of decomposition can be considered as long as the properties associated with $\mathbf{U}$ and $\mathbf{V}$ are preserved. For example, Bivariate Moment Decomposition (BMD), which is especially applicable for ordinal categories, may be considered in the correspondence analysis of a two-way contingency table; see, for example, Beh (1997,

1998) and Lombardo et al. (2007). Alteratively, a Hybrid Decomposition (HD) utilising features of SVD and BMD may be considered; Beh (2001, 2008) and Lombardo et al. (2011).

References

Agresti A (2002) Categorical data analysis. Wiley, London

Aitchison J, Greenacre M (2002) Biplots of compositional data. Appl Stat 51:375–392

Andrews D (1972) Plots of high dimensional data. Biometrics 28:125–136

Beh EJ (1997) Simple correspondence analysis of ordinal cross-classifications using orthogonal polynomials. Biom J 39:589–613

Beh EJ (1998) A comparative study of scores for correspondence analysis with ordered categories. Biom J 40:413–429

Beh EJ (2001) Partitioning Pearson's chi-squared statistic for singly ordered two-way contingency tables. Aust N Z J Stat 43:327–333

Beh EJ (2004) Simple correspondence analysis: a bibliographic review. Int Stat Rev 72:257–284

Beh EJ (2008) Simple correspondence analysis of nominal-ordinal contingency tables. J Appl Math Decis Sci 2008:218140

Beh EJ (2010) Elliptical confidence regions for simple correspondence analysis. J Stat Plan Inference 140:2582–2588

Beh EJ (2012) Simple correspondence analysis using adjusted residuals. J Stat Plan Inference 142:965–973

Beh EJ, Lombardo R (2012) A genealogy of correspondence analysis. Aust N Z J Stat 54:137–168

Beh EJ, Lombardo R (2014) Correspondence analysis: theory, practice and new strategies. Wiley, London

Beh EJ, Smith DR (2011) Real world occupational epidemiology, part 1: a visual interpretation of statistical significance. Arch Environ Occup Health 66:119–123

Best DJ (1995) Consumer data: statistical tests for differences in dispersion. Food Qual Prefer 6:221–225

Best DJ, Rayner JCW, O'Sullivan MG (1999) Product maps for sensory ranking and categorical data. In: Watson AJ, Bell GA (eds) Tastes and aromas, the chemical senses in science and industry. Wiley, New York, pp 114–119

Best DJ, Rayner JCW, O'Sullivan MG (2000) Product maps for consumer categorical data. Food Qual Prefer 11:91–97

Chernoff H (1973) The use of faces to represent points in k-dimensional space graphically. J Am Stat Assoc 68:361–368

Choulakian V (1988) Exploratory analysis of contingency tables by loglinear formulation and generalizations of correspondence analysis. Psychometrika 53:235–250

D'Ambra L, Lauro NC (1989) Non symmetrical analysis of three way contingency tables. In: Coppi R, Bolasco S (eds) Multiway data analysis. North-Holland, Amsterdam, pp 301–315

D'Ambra L, Lauro NC (1992) Non-symmetrical exploratory data analysis. Stat Appl 4:511–529

De Leeuw J (1993) Some generalizations of correspondence analysis. In: Cuadras CM, Rao CR (eds) Multivariate analysis: future directions 2. North-Holland, Amsterdam, pp 359–375

Escoufier Y (1988) Assessing the number of axes that should be considered in correspondence analysis. In: Hayashi C, Diday E, Jambu M (eds) Recent developments in clustering and data analysis. Academic Press, Cambridge, pp 231–240

Freeman MF, Tukey JW (1950) Transformations related to the angular square root. Ann Math Stat 21:607–611

Friendly M (1994) Mosaic displays for multi-way contingency tables. J Am Stat Assoc 89:190–200

Friendly M (1999) Extending mosaic displays: marginal, conditional, and partial views of categorical data. J Comput Graph Stat 8:373–395

Gabriel KR (1971) The biplot graphic display of matrices with application to principal component analysis. Biometrika 58:453–467

Gabriel KR, Orodoff CL (1990) Biplots in biomedical research. Stat Med 9:469–485

Gilula Z, Krieger M, Ritov Y (1988) Ordinal association in contingency tables: some interpretative aspects. J Am Stat Assoc 83:540–545

Goodman LA (1986) Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. Int Stat Rev 54:243–309

Goodman LA (1996) A single general method for the analysis of cross-classified data: reconciliation and synthesis of some methods of Pearson, Yule, Fisher, and also some methods of correspondence analysis and association analysis. J Am Stat Assoc 91:408–428

Goodman LA, Kruskal WH (1954) Measures of association for cross-classifications. J Am Stat Assoc 49:732–764

Gower J, Lubbe S, le Roux N (2011) Understanding biplots. Wiley, London

Gower JC, le Roux NJ, Gardner-Lubbe S (2014) Biplots: quantitative data. WIREs Comput Stat 7:42–62

Grassi M, Visentin S (1994) Correspondence analysis applied to grouped cohort data. Stat Med 13:2407–2425

Greenacre MJ (1984) Theory and application of correspondence analysis. Academic Press, London

Greenacre MJ (1990) Some limitations of multiple correspondence analysis. Comput Stat Quaterly 3:249–256

Greenacre MJ (1993) Biplots in correspondence analysis. J Appl Stat 20:251–269

Greenacre M (2010a) Biplots in practice. Fundacion BBVA, Madrid

Greenacre M (2010b) Correspondence analysis. WIRE's Comput Stat 2:613–619

Greenacre M, Blasius J (eds) (2006) Multiple correspondence analysis and related methods. Chapman & Hall/CRC, Boca Raton

Haberman SJ (1973) The analysis of residuals in cross-classified tables. Biometrics 29:205–220

Kroonenberg PM, Lombardo R (1998) Nonsymmetric correspondence analysis: a tutorial. Kwant Methoden 58:57–83

Linting M, Meulman JJ, Groenen PJF, van der Kooij AJ (2007) Stability of nonlinear principal components analysis: an empirical study using the balanced bootstrap. Psychol Methods 12:359–379

Lombardo R, Ringrose T (2012) Bootstrap confidence regions in non-symmetrical correspondence analysis. Electron J Appl Stat Anal 5:413–417

Lombardo R, Carlier A, D'Ambra L (1996) Nonsymmetric correspondence analysis for three-way contingency tables. Methodologica 4:59–80

Lombardo R, Beh EJ, D'Ambra L (2007) Non-symmetric correspondence analysis with ordinal variables. Comput Stat Data Anal 52:566–577

Lombardo R, Beh EJ, D'Ambra A (2011) Studying the dependence between ordinal-nominal categorical variables via orthogonal polynomials. J Appl Stat 38:2119–2132

Lombardo R, Beh EJ, Kroonenberg PM (2016) Modelling trends in ordered correspondence analysis using orthogonal polynomials. Psychometrika 81:325–349

Lorenzo-Seva U, van de Velden M, Kiers HAL (2009) Oblique rotation in correspondence analysis: a step forward in the search for the simplest interpretation. Br J Math Stat Psychol 62:583–600

Markus MT (1994) Bootstrap confidence regions in non-linear multivariate analysis. DSWO Press, Leiden

Nishisato S (2007) Multidimensional nonlinear descriptive analysis. Chapman & Hall/CRC, London

Ringrose TJ (2012) Bootstrap confidence regions for correspondence analysis. J Stat Comput Simul 82:1397–1413

Selikoff IJ (1981) Household risks with inorganic fibers. Bull N Y Acad Med 57:947–961

Upton GJG (2000) Cobweb diagrams for multiway contingency tables. The Statistician 49:79–85

Upton GJG (2016) Categorical data analysis by example. Wiley, Hoboken

Van de Velden M, Kiers HAL (2005) Rotation in correspondence analysis. J Classif 22:251–271

Yanai H (1986) Some generalizations of correspondence analysis in terms of projection operators. In: Diday E, Escoufier Y, Lebart L, Pagés JP, Schektman Y, Thomassone R (eds) Data analysis and informatics IV. North-Holland, Amsterdam, pp 193–207