

Article

Visualising Departures from Symmetry and Bowker's X^2 Statistic

Eric J. Beh ^{1,2,*}  and Rosaria Lombardo ^{3,†} 

¹ National Institute for Applied Statistics Research Australia (NIASRA), University of Wollongong, Wollongong, NSW 2522, Australia

² Department of Statistics and Actuarial Science, Stellenbosch University, Stellenbosch 7602, South Africa

³ Department of Economics, University of Campania, 810423 Capua, CE, Italy; rosaria.lombardo@unicampania.it

* Correspondence: ericb@uow.edu.au

† These authors contributed equally to this work.

Abstract: Sometimes, the same categorical variable is studied over different time periods or across different cohorts at the same time. One may consider, for example, a study of voting behaviour of different age groups across different elections, or the study of the same variable exposed to a child and a parent. For such studies, it is interesting to investigate how similar, or different, the variable is between the two time points or cohorts and so a study of the departure from symmetry of the variable is important. In this paper, we present a method of visualising any departures from symmetry using correspondence analysis. Typically, correspondence analysis uses Pearson's chi-squared statistic as the foundation for all of its numerical and visual features. In the case of studying the symmetry of a variable, Bowker's chi-squared statistic, presented in 1948, provides a simple numerical means of assessing symmetry. Therefore, this paper shall discuss how a correspondence analysis can be performed to study the symmetry (or lack thereof) of a categorical variable when Bowker's statistic is considered. The technique presented here provides an extension to the approach developed by Michael Greenacre in 2000.

Keywords: correspondence analysis; Bowker's chi-squared statistic; singular value decomposition; skew-symmetric matrix



Citation: Beh, E.J.; Lombardo, R. Visualising Departures from Symmetry and Bowker's X^2 Statistic. *Symmetry* **2022**, *14*, 1103. <https://doi.org/10.3390/sym14061103>

Academic Editor: Kouji Tahata

Received: 8 April 2022

Accepted: 12 May 2022

Published: 27 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Studying the departure from symmetry between categorical variables that are cross-classified to form a contingency table has been a topic of much discussion over the past few decades. There is wealth of literature that is now available that examines the measuring, modelling and application of symmetric categorical variables including, but not confined to, the contributions of Agresti [1] (Section 10.4), Anderson [2] (Section 10.2), Bove [3], De Falguerolles and van der Heijden [4], Iki, Yamamoto and Tomizawa [5], Tomizawa [6], Yamamoto [7] and Yamamoto, Shimada and Tomizawa [8]. Most recently, Arellano-Valle, Contreras-Reyes and Stehlik [9] and Nishisato [10] provides a wide ranging discussion on symmetry from the perspective of quantification theory. Such techniques are generally applied to an $S \times S$ contingency table, denoted here by \mathbf{N} , where a categorical variable may be examined over, for example, two time periods or between two cohorts. When interest is on assessing departures from *complete* or *perfect* symmetry in \mathbf{N} , further information on the symmetric (or lack thereof) nature of the variables may be gained by partitioning \mathbf{N} so that

$$\mathbf{N} = \mathbf{Y} + \mathbf{K} = \frac{1}{2}(\mathbf{N} + \mathbf{N}^T) + \frac{1}{2}(\mathbf{N} - \mathbf{N}^T) \quad (1)$$

where \mathbf{Y} is the matrix reflecting the symmetric part of \mathbf{N} while \mathbf{K} is the skew-symmetric part of the matrix. This partition was considered by Bove [3], Constantine and Gower [11]

(Section 3), Gower [12] and Greenacre [13] for visualising row and column categories that deviate from the null hypothesis of perfect symmetry. Greenacre [13] shows how (1) can be used when performing a correspondence analysis on \mathbf{N} . His approach yields two visual summaries—one reflecting the symmetric structure (by performing a correspondence analysis on \mathbf{Y}) and another that visualises that part of the structure that is not symmetric (through a correspondence analysis of \mathbf{K}).

In this paper, we also consider how one may perform a correspondence analysis on a contingency table formed from the cross-classification of symmetric variables. Our strategy is different, but not completely independent, to that of Greenacre's [13] method. The similarities of our approach and those of Greenacre [13] include basing our analysis on a weighted version of \mathbf{K} that yields pairs of equivalent principal inertia values. However, the key difference of our technique when compared with Greenacre's [13] is that we show how the approach outlined below can be performed when the underlying measure of the departure from perfect symmetry is Bowker's [14] chi-squared statistic. By doing so, we show, for example, that there exists a link between the principal coordinates and singular values that are obtained and Bowker's statistic.

For our discussion, this paper consists of seven further sections. Section 2 provides an overview of correspondence analysis which is typically used for visually exploring departures from complete independence between the categorical variables. In this section, we describe the application of singular value decomposition (SVD) to a transformation of a standard two-way contingency table, we define the principal coordinates needed to construct the visual summary and describe other key features. We then turn our attention to presenting an overview of the test of symmetry between the two categorical variables and the role that Bowker's chi-squared statistic plays in such a test; see Section 3. Section 4 then shows how Bowker's statistic can be used as the core global measure for assessing the variable's departure from symmetry when providing a visual summary of such sources of departure using correspondence analysis. Section 5 illustrates some features of correspondence analysis when applied to Bowker's statistic. Two examples are then presented that demonstrate the applicability of the technique. Section 6 applies this method of correspondence analysis to an artificially constructed 4×4 contingency table that exhibits perfect symmetry when a constant $C = 0$ is added to the (2, 1)th cell frequency. Changes in C then provide a means of visualising the extent to which the rows and columns deviate from what is expected when the two variables are perfectly symmetric. Our second example, studied in Section 7, considers the data presented in Grover and Srinivasan [15], which looks at differences in the purchase of five brands of decaffeinated coffee across two time periods. Section 8 provides some final comments on the technique, including possible extensions and emendations that may be considered for future research. Three appendices are also included. The first two appendices derive the singular values for a contingency table of size 2×2 and 3×3 (both of which always yields an optimal display that consists of two dimensions). The third appendix provides a description of the R function `bowkerca.exe()` that performs the necessary calculations of the approach described in this paper.

2. On the Classical Approach

Before we outline our approach to the correspondence analysis of a contingency table using Bowker's chi-squared statistic it is worth providing a broad overview of the classical technique; one may also consider, for example, Beh and Lombardo [16] for a detailed historical, methodological, practical and computational discussion of a range of correspondence analysis methods. To do so, we consider a contingency table, \mathbf{N} of size $I \times J$ so that there are I row categories and J column categories. The (i, j) th cell frequency of \mathbf{N} is denoted by n_{ij} so that the total sample size is $n = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$. Denote the matrix of relative proportions by \mathbf{P} where the (i, j) th value is $p_{ij} = n_{ij}/n$ so that the sum of these proportions across all IJ cells of the table is 1. We also define the i th row marginal proportion by $p_{i\bullet} = \sum_{j=1}^J p_{ij}$ so that it is the i th element of the vector \mathbf{r} and the (i, i) th element of the diagonal matrix \mathbf{D}_R . Similarly, the j th column marginal proportion

is denoted by $p_{\bullet j} = \sum_{i=1}^I p_{ij}$, so that it is the j th element of the vector \mathbf{c} and the (j, j) th element of the diagonal matrix \mathbf{D}_C .

To test whether the observed set of proportions, p_{ij} , differs from what is expected under some model with an expected value of \hat{p}_{ij} , then Pearson's chi-squared statistic for such a test takes the form

$$X^2 = n \sum_{i=1}^S \sum_{j=1}^S \frac{(p_{ij} - \hat{p}_{ij})^2}{\hat{p}_{ij}}. \quad (2)$$

For example, under the hypothesis of complete independence $\hat{p}_{ij} = p_{i\bullet}p_{\bullet j}$ so that (2) becomes

$$X_I^2 = n \sum_{i=1}^I \sum_{j=1}^J r_{ij}^2 \quad (3)$$

where

$$r_{ij} = \frac{p_{ij} - p_{i\bullet}p_{\bullet j}}{\sqrt{p_{i\bullet}p_{\bullet j}}}$$

is the (i, j) th standardised (Pearson) residual and X_I^2 is a chi-squared random variable with $(I - 1)(J - 1)$ degrees of freedom. If there is a statistically significant association between the row and column variables this association can be visualised using correspondence analysis. This is achieved by first performing a SVD on the matrix of standardised residuals, such that

$$\mathbf{D}_R^{-1/2}(\mathbf{P} - \mathbf{rc}^T)\mathbf{D}_C^{-1/2} = \mathbf{AD}_\lambda\mathbf{B}^T.$$

Here, \mathbf{A} and \mathbf{B} are the column matrices containing the left and right singular vectors, respectively, and are constrained such that $\mathbf{A}^T\mathbf{A} = \mathbf{I}_M$ and $\mathbf{B}^T\mathbf{B} = \mathbf{I}_M$, where $M = \min(I, J) - 1$; here \mathbf{I}_M is an $M \times M$ identity matrix. The (m, m) th element of the diagonal matrix \mathbf{D}_λ is the m th singular value and these values arranged in descending order so that $1 > \lambda_1 > \dots > \lambda_M > 0$.

A visual depiction of the association between the row and column variables can be made by simultaneously projecting the row principal coordinates and the column principal coordinates onto the same low-dimensional space which optimally consists of M dimensions; such a display is commonly referred to as a correspondence plot and typically consists of the first two dimensions (for ease of visualisation). The matrix of row and column principal coordinates are defined by

$$\begin{aligned} \mathbf{F} &= \mathbf{D}_R^{-1/2}\mathbf{AD}_\lambda \\ \mathbf{G} &= \mathbf{D}_C^{-1/2}\mathbf{BD}_\lambda \end{aligned}$$

respectively.

Since (3) shows that Pearson's chi-squared statistic is linearly related to the sample size, n , correspondence analysis uses X^2/n as the measure of association which is termed the total inertia of the contingency table. This measure is directly related to the principal coordinates and singular values, such that

$$\begin{aligned} \frac{X^2}{n} &= \text{trace}(\mathbf{D}_\lambda^2) \\ &= \text{trace}(\mathbf{F}^T\mathbf{D}_R\mathbf{F}) \\ &= \text{trace}(\mathbf{G}^T\mathbf{D}_C\mathbf{G}). \end{aligned}$$

Therefore, points located at a great distance from the origin provide a visual indication of the importance that a category plays in the association structure of the variables, while the origin is the position of all of the row and column principal coordinates if there is complete independence between the variables.

We shall not provide a comprehensive account of all of the features, and related methods, of correspondence analysis. Instead, the interested reader is directed to Beh and Lombardo [16], for example, for more information.

3. On Studying the Symmetry of a Categorical Variable

Suppose we now wish to study the departure from complete symmetry of two categorical variables that share a similar structure. Let S be the number of categories contained in this variable so that the contingency table is now of size $S \times S$. Then, a test of symmetry of the variable may be undertaken by defining the null hypothesis by

$$H_0 : p_{ij} = p_{ji}; \forall i, j.$$

Sometimes, a study of symmetry in a contingency table may be undertaken by assessing the *marginal homogeneity* of the table where the null hypothesis is

$$H_0 : p_{i\bullet} = p_{\bullet j}; \forall i, j$$

but we shall say very little on this issue in the following sections.

When assessing whether there is any evidence of symmetry between the row and column variable of a contingency table, the most appropriate choice of \hat{p}_{ij} is

$$\hat{p}_{ij} = \frac{p_{ij} + p_{ji}}{2}; \quad (4)$$

see, for example, Agresti [1] (p. 424) and Anderson [2] (p. 321). Substituting (4) into (2), and denoting the resulting statistic by X_S^2 , yields

$$X_S^2 = \frac{n}{2} \sum_{i=1}^S \sum_{j=1}^S \frac{(p_{ij} - p_{ji})^2}{p_{ij} + p_{ji}} \quad (5)$$

$$= n \sum_{i>j} \frac{(p_{ij} - p_{ji})^2}{p_{ij} + p_{ji}} \quad (6)$$

and is the chi-squared statistic derived by Bowker [14] for testing the symmetry between a row and column variable of a contingency table and has $S(S-1)/2$ degrees of freedom. When $S = 2$ (6) simplifies to McNemar's [17] statistic for testing the symmetry of two cross-classified dichotomous variables.

The simplicity of using (6) has gained wide appeal and was considered in the classic texts of Agresti [1] (p. 424), Bishop, Fienberg and Holland [18] (p. 283), Lancaster [19] (p. 236) and Plackett [20] (p. 59). However, Agresti [1] (p. 425) does point out that "it rarely fits well. When the marginal distributions differ substantially, it fits poorly". The concern here about the poor fit when the margins of the contingency table differ substantially assumes that one is also interested in testing the independence between the categories. This is because when one assumes independence in the context of symmetry (4) becomes

$$\hat{p}_{ij} = \frac{p_{i\bullet} p_{\bullet j} + p_{j\bullet} p_{\bullet i}}{2}$$

and requires that $p_{i\bullet} = p_{\bullet i}$ for all $i = 1, 2, \dots, S$, which we need not impose for testing symmetry. Perhaps what helps to clarify this point is that Lancaster [19] (p. 237) says that "more difficulties arise if... it is desired to test the homogeneity of the margins since p_{ij} may not be equal to $p_{i\bullet} p_{\bullet j}$ ". Since we are not concerned with testing for marginal homogeneity in this paper this "difficulty" is of no concern to us. We are also not greatly concerned with the inferential aspects of the statistic since (6) is used as a numerical basis on which correspondence analysis lies; in doing so, we are assuming, or have *a priori* tested and confirmed, that there is a departure from symmetry and we wish only to visualise the potential sources of this departure.

One may also consider the log-likelihood ratio statistic

$$\begin{aligned} G_S^2 &= 2n \sum_{i=1}^I \sum_{j=1}^I p_{ij} \ln \left(\frac{p_{ij}}{\hat{p}_{ij}} \right) \\ &= 2n \sum_{i \neq j} p_{ij} \ln \left(\frac{2p_{ij}}{p_{ij} + p_{ji}} \right) \end{aligned}$$

as an alternative to X_S^2 ; see Bishop, Fienberg and Holland [18] (p. 283). We shall not discuss how this statistic can be used for studying departures from symmetry using correspondence analysis. However, the log-likelihood statistic, like Pearson's statistic of (3), is a special case of the Cressie–Read family of divergence statistics [21] and Beh and Lombardo [22] provide an overview of how one may perform correspondence analysis using this divergence statistic. Future research can certainly be undertaken to study the role of G_S^2 in correspondence analysis when studying the symmetry of a categorical variable.

4. On Bowker's Residuals and Departures from Symmetry

Let \mathbf{S} be the matrix of the *Bowker residuals* where the (i, j) th element is

$$s_{ij} = \frac{1}{\sqrt{2}} \frac{p_{ij} - p_{ji}}{\sqrt{p_{ij} + p_{ji}}} \quad (7)$$

so that Bowker's statistic, X_S^2 defined by (5), can be expressed as

$$\begin{aligned} X_S^2 &= n \sum_{i=1}^S \sum_{j=1}^S s_{ij}^2 \\ &= n \text{trace}(\mathbf{S}^T \mathbf{S}) \\ &= n \text{trace}(\mathbf{S} \mathbf{S}^T). \end{aligned} \quad (8)$$

A feature of the matrix \mathbf{S} (and \mathbf{K} in (1)) is that it is an *anti-symmetric*, or skew-symmetric, matrix so that $\mathbf{S}^T = -\mathbf{S}$. Therefore, the left and right singular vectors, and the singular values, of the matrix of the *Bowker residuals*, \mathbf{S} , can be obtained from the eigen-decomposition of $\mathbf{S}^T \mathbf{S}$ or, equivalently, of $-\mathbf{S}^2$. If S is odd then there will always be a zero eigen-value and $S - 1$ positive eigen-values; see Ward and Gray [23] and Gower [12] (p. 113). Ward and Gray [23] also present an algorithm that can perform the necessary eigen-decomposition and in Appendices A and B we provide a derivation that leads to closed-form solutions of the eigen-values of a 2×2 and 3×3 matrix of the *Bowker residuals*. Constantine and Gower [11] note that, since $\mathbf{S}^T \mathbf{S} = -\mathbf{S}^2$, there will be pairs of identical eigen-values. This is a feature that is also demonstrated in the appendices and described in Sections 6 and 7.

Suppose we denote the (i, j) th element of \mathbf{Y} and \mathbf{K} from (1) by y_{ij} and k_{ij} , respectively. Then, the (i, j) th Bowker residual—see (7)—can be alternatively expressed as

$$s_{ij} = \frac{1}{2} (p_{ij} - p_{ji}) \div \sqrt{\frac{1}{2} (p_{ij} + p_{ji})} = \frac{k_{ij}}{\sqrt{y_{ij}}}. \quad (9)$$

Therefore, Bowker's residuals do assess for each cell of the contingency table where departures from symmetry exist but they do so relative to the amount of symmetry that exists between the two variables. This is in contrast to Greenacre's [13] approach which considers a residual where k_{ij} is divided by the mean of the row and column marginal proportions; for the i th such proportion this is $(p_{i\bullet} + p_{\bullet i})/2$ for $i = 1, 2, \dots, S$.

In addition, since the numerator of \mathbf{S} can be expressed in terms of \mathbf{K} this ensures that Bowker's residuals are centred at zero under perfect symmetry (a property shared with Greenacre's [13] approach). This is important since it means that the principal coordinates that we derive in Section 5.3 are centred at the origin of the optimal correspondence plot.

5. Correspondence Analysis and Bowker's Statistic

5.1. On the SVD of the Matrix of the Bowker Residuals

Visually detecting departures from symmetry can be undertaken using correspondence analysis. This can be achieved by first applying a SVD to \mathbf{S} such that

$$\mathbf{S} = \mathbf{A}\mathbf{D}_\lambda\mathbf{B}^T. \quad (10)$$

Here, the $S \times M$ column matrix \mathbf{A} contains the left singular vectors of \mathbf{S} and have the property $\mathbf{A}^T\mathbf{A} = \mathbf{I}_M$. Similarly, \mathbf{B} is a $S \times M$ column matrix of the right singular vectors such that $\mathbf{B}^T\mathbf{B} = \mathbf{I}_M$ of \mathbf{S} . While $M = \min(S, S) - 1 = S - 1$ for the classical approach to correspondence analysis, this is not the case for the analysis of the matrix of the Bowker residuals. Since \mathbf{S} is a skew-symmetric matrix there will be S singular values when S is even and $S - 1$ such values when S is odd. Therefore,

$$M = \begin{cases} S, & \text{if } S \text{ is even} \\ S - 1, & \text{if } S \text{ is odd} \end{cases}.$$

Gower [12] showed that since \mathbf{S} is skew-symmetric then the SVD of the matrix, (10), is equivalent to

$$\mathbf{S} = \mathbf{A}\mathbf{D}_\lambda\mathbf{J}_M\mathbf{A}^T \quad (11)$$

so that

$$\mathbf{B} = \mathbf{A}\mathbf{J}_M^T. \quad (12)$$

Here \mathbf{J}_M is a $M \times M$ block diagonal and orthogonal skew-symmetric matrix so that

$$\mathbf{J}_M^T\mathbf{J}_M = \mathbf{J}_M\mathbf{J}_M^T = \mathbf{I}_M \quad (13)$$

where \mathbf{I}_M is an $M \times M$ identity matrix. When $M = 2$,

$$\mathbf{J}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

while

$$\mathbf{J}_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

If S is odd then the (S, S) th element is 1 and the rest of its row/column consists of zeros so that

$$\mathbf{J}_S = \begin{pmatrix} \mathbf{J}_{S-1} & \mathbf{0}_{S-1}^T \\ \mathbf{0}_{S-1} & 1 \end{pmatrix}$$

where $\mathbf{0}_{S-1}$ is a zero vector of length $S - 1$. For example,

$$\mathbf{J}_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The $M \times M$ diagonal matrix \mathbf{D}_λ contains the singular values of \mathbf{S} and are arranged in consecutive pairs of values so that $1 > \lambda_1 = \lambda_2 > \lambda_3 = \lambda_4 > \dots \geq 0$. This feature will be discussed in the next section, but it is worth noting that the SVD presented in (11) makes use of the *Murnaghan canonical form* of \mathbf{S} ; see Murnaghan and Wintner [24] and Paardekooper [25] for further information on this decomposition.

5.2. The Total Inertia

Quantifying the departure from symmetry can be undertaken by calculating the total inertia of \mathbf{S} which is just Bowker's statistic divided by the sample size, X_S^2/n . When \mathbf{B} , for example, is of full rank, then $\mathbf{B}\mathbf{B}^T = \mathbf{B}^T\mathbf{B}$. Thus, using (8), the total inertia can be expressed as the sum-of-squares of the singular values such that

$$\begin{aligned}\phi_S^2 &= \frac{X_S^2}{n} \\ &= \text{trace}\left(\left(\mathbf{A}\mathbf{D}_\lambda\mathbf{B}^T\right)^T\left(\mathbf{A}\mathbf{D}_\lambda\mathbf{B}^T\right)\right) \\ &= \text{trace}\left(\mathbf{B}\mathbf{D}_\lambda^2\mathbf{B}^T\right) \\ &= \text{trace}\left(\mathbf{D}_\lambda^2\right).\end{aligned}$$

When visualising departures from symmetry, one may construct a correspondence plot consisting of at most M dimensions where the principal inertia (or weight) of the m th dimension is λ_m^2 . Therefore, expressing the total inertia in terms of the matrix \mathbf{D}_λ^2 (or the elements λ_m^2) provides a means of determining, for each dimension, the percentage of the total departure from symmetry that exists between the row and column variables. Such a percentage is calculated by

$$100 \times \frac{\lambda_m^2}{\phi_S^2}.$$

Since $\lambda_1 = \lambda_2$ is the largest pair of singular values, the first and second dimensions visually provide the best (and an equivalent quality) depiction of any departure from symmetry that exists in \mathbf{N} . For example, when $S = 5$, say, the first and second dimensions will always provide the same (and most) amount of information to the visual display, while the third and fourth dimensions will be equally weighted and display less of the departure. These four dimensions will provide an optimal display of the departure from symmetry since the fifth dimension will have a zero principal inertia value.

5.3. The Principal Coordinates

When calculating Bowker's statistic, only the cell proportions p_{ij} and p_{ji} are used, not the marginal proportions; see (5) and equivalently (6). The cloud-of-points generated for the rows and columns can be obtained by aggregating these cell proportions across the row and column variables, yielding two spaces that have the same metric; this metric being $\hat{\mathbf{D}} = (\mathbf{D}_R + \mathbf{D}_C^T)/2$. Such a property is consistent, but slightly different, to the traditional approach to correspondence analysis which assesses departures from independence, not symmetry, and so assumes that the row space and column space have a metric based on the aggregation of p_{ij} across the two variables. That is, the row space has the metric \mathbf{D}_R while the column metric is \mathbf{D}_C . Thus, the metric $\hat{\mathbf{D}}$ is the average of the row and column spaces and is consistent with the metric adopted by Greenacre [13].

To obtain a visual depiction of the departures from symmetry one may simultaneously plot the row and column principal coordinates which are defined in terms of the above matrices by

$$\mathbf{F} = \hat{\mathbf{D}}^{-1/2}\mathbf{A}\mathbf{D}_\lambda \quad (14)$$

$$\mathbf{G} = \hat{\mathbf{D}}^{-1/2}\mathbf{B}\mathbf{D}_\lambda \quad (15)$$

respectively.

5.4. On the Origin and Transition Formulae

The total inertia of the two-way contingency table can be expressed in terms of the principal coordinates defined by (14) and (15). For example, for the row coordinates,

$$\begin{aligned}\text{trace}(\mathbf{F}^T \hat{\mathbf{D}} \mathbf{F}) &= \text{trace}\left(\left(\hat{\mathbf{D}}^{-1/2} \mathbf{A} \mathbf{D}_\lambda\right)^T \hat{\mathbf{D}} \left(\hat{\mathbf{D}}^{-1/2} \mathbf{A} \mathbf{D}_\lambda\right)\right) \\ &= \text{trace}\left(\mathbf{D}_\lambda^T \mathbf{A}^T \mathbf{A} \mathbf{D}_\lambda\right) \\ &= \text{trace}\left(\mathbf{D}_\lambda^2\right) \\ &= \frac{X_S^2}{n}.\end{aligned}$$

This shows that the origin coincides with the location of all row coordinates when there is perfect symmetry between variables of the contingency table. By following a similar derivation, it can also be shown that

$$\frac{X_S^2}{n} = \text{trace}\left(\mathbf{G}^T \hat{\mathbf{D}} \mathbf{G}\right).$$

These expressions of the total inertia also show that points located far from the origin identify categories that deviate from what is expected if there was perfect symmetry.

Alternative expressions of the matrices of row and column principal coordinates can also be obtained. For example, suppose we consider (14). Then, by post-multiplying both sides by \mathbf{B}^T and simplifying we get

$$\mathbf{F} = \hat{\mathbf{D}}^{-1/2} \mathbf{S} \mathbf{B}.$$

Thus, the (i, m) 'th element of \mathbf{F} is

$$\begin{aligned}f_{im} &= \frac{\sqrt{2}}{\sqrt{p_{i\bullet} + p_{\bullet i}}} \sum_{j=1}^J s_{ij} b_{jm} \\ &= \frac{\sqrt{2}}{\sqrt{p_{i\bullet} + p_{\bullet i}}} \sum_{j=1}^J \frac{1}{\sqrt{2}} \frac{p_{ij} - p_{ji}}{\sqrt{p_{ij} + p_{ji}}} b_{jm} \\ &= \sum_{j=1}^J \frac{p_{ij} - p_{ji}}{\sqrt{(p_{i\bullet} + p_{\bullet i})(p_{ij} + p_{ji})}} b_{jm}.\end{aligned}$$

Therefore, a principal coordinate will lie at the origin of the correspondence plot if there is perfect symmetry between the i th row and i th column; that is, when $p_{ij} = p_{ji}$ for all $j = 1, 2, \dots, S$, thus verifying expressing the total inertia in terms of the matrix of row principal coordinates. Any departure from symmetry of the i th row from the i th column will result in the i th principal coordinate moving away from the origin. This implies that the i th column category will also move away from the origin but it will do so in a different direction from the row category. Such a feature can be verified by showing that the row and column principal coordinates are linked through the following transition formulae. By substituting (12) into (15) the matrix of column principal coordinates can be alternatively expressed as

$$\mathbf{G} = \hat{\mathbf{D}}^{-1/2} \mathbf{A} \mathbf{J}_M^T \mathbf{D}_\lambda = \hat{\mathbf{D}}^{-1/2} \mathbf{A} \mathbf{D}_\lambda \mathbf{J}_M^T.$$

Thus, the matrix of column principal coordinates can be expressed in terms of the matrix of row principal coordinates such that

$$\mathbf{G} = \mathbf{F} \mathbf{J}_M^T. \quad (16)$$

Similarly, post-multiplying both sides of (16) by \mathbf{J}_M and using (13) leads to

$$\mathbf{F} = \mathbf{G}\mathbf{J}_M. \quad (17)$$

showing how the matrix of row principal coordinates can be expressed in terms of the matrix of column principal coordinates. Thus, for the i th row and column categories, $(f_{i1}, f_{i2}) = (-g_{i2}, g_{i1})$ showing that departures from symmetry between these two categories will position their principal coordinates on opposite sides of the correspondence plot, unless there is perfect symmetry in which case both will lie at the origin.

5.5. Intra-Variable Distances

Suppose we are interested in the distance between the i th and i' th row principal coordinates in a correspondence plot. It can be shown that the squared Euclidean distance between f_{im} and $f_{i'm}$ is

$$d_1^2(i, i') = \sum_{j=1}^S c_{ij} \left[\frac{p_{ij} + p_{i'j}}{2} - \frac{p_{ji} + p_{j'i'}}{2} \right]^2 \quad (18)$$

$$= \sum_{j=1}^S c_{ij} \left[\frac{p_{ij} - p_{ji}}{2} + \frac{p_{i'j} - p_{j'i'}}{2} \right]^2 \quad (19)$$

where

$$c_{ij} = \frac{p_{\bullet j} + p_{j\bullet}}{\sqrt{(p_{ij} + p_{ji})(p_{i'j} + p_{j'i'})}}$$

Thus, if $p_{ij} + p_{i'j} = p_{ji} + p_{j'i'}$ or, equivalently, $\hat{p}_{ij} = \hat{p}_{i'j}$ for all $j = 1, 2, \dots, S$ then the i th and i' th row principal coordinates will be located in the same position in the optimal correspondence plot. Such a feature arises when the expected cell frequency (under perfect symmetry) is the same for the i th and i' th rows. Alternatively, $f_{im} = f_{i'm}$ for all $m = 1, 2, \dots, S$ when $p_{ij} - p_{ji} = p_{i'j} - p_{j'i'}$. This feature does not imply that there must be perfect symmetry between the i th and i' th row categories, but it does arise when any departure from perfect symmetry is the same for these categories.

6. Example 1 Artificial Data

6.1. The Data

To examine how to perform a correspondence analysis on a contingency table using Bowker's statistic (6), suppose we consider the artificial data presented in Table 1. Here, $C \geq 0$ in the (2, 1)th cell is a non-negative integer ensuring that Table 1 maintains the features of a contingency table. For this table, Bowker's statistic is

$$\begin{aligned} X_5^2 &= \sum_{i>j}^4 \frac{(n_{ij} - n_{ji})^2}{n_{ij} + n_{ji}} \\ &= \frac{(n_{21} - n_{12})^2}{n_{21} + n_{12}} + \frac{(n_{31} - n_{13})^2}{n_{31} + n_{13}} + \frac{(n_{41} - n_{14})^2}{n_{41} + n_{14}} \end{aligned} \quad (20)$$

$$\begin{aligned} &+ \frac{(n_{32} - n_{23})^2}{n_{32} + n_{23}} + \frac{(n_{42} - n_{24})^2}{n_{42} + n_{24}} + \frac{(n_{43} - n_{34})^2}{n_{43} + n_{34}} \\ &= \frac{C^2}{40 + C} \end{aligned} \quad (21)$$

and is a chi-squared random variable with $4 \times (4 - 1) / 2 = 6$ degrees of freedom. Therefore, (21) shows that when $C = 0$, Bowker's chi-squared statistic is zero and Table 1 is perfectly symmetric.

Table 1. A near-symmetric artificial contingency table where C is a non-negative integer.

	C1	C2	C3	C4
R1	10	20	30	40
R2	$20 + C$	50	60	70
R3	30	60	20	40
R4	40	70	40	80

6.2. Preliminary Examination of the Departure from Symmetry

Departures from symmetry for Table 1 can be assessed by determining the minimum value of C so that $X_{\zeta}^2 > \chi_{\alpha}^2(6)$. That is, when

$$\frac{C^2}{40 + C} > \chi_{\alpha}^2(6).$$

From this result, one may obtain the following quadratic equation (in terms of C)

$$C^2 - \chi_{\alpha}^2(6) \times C - 40\chi_{\alpha}^2(6) > 0$$

which only has one valid solution

$$C = \frac{\chi_{\alpha}^2(6) + \sqrt{(\chi_{\alpha}^2(6))^2 + 160\chi_{\alpha}^2(6)}}{2}$$

that ensures that C is non-negative. For example, when testing symmetry using $\alpha = 0.05$, $\chi_{0.05}^2(6) = 12.59$ so that C must exceed 29.59 for the departure from symmetry to be statistically significant. For practical purposes, C must be an integer of at least 30 to ensure that the (2, 1)th cell remains a positive integer. Hence, this cell frequency must exceed 50 for there to be a statistically significant lack of symmetry in Table 1. Thus, to demonstrate how the correspondence analysis technique described in Section 5 can be applied to Table 1 we shall be considering the following values of C ; 50, 75, 100 and 150. These values of C give a Bowker statistic of 27.78, 48.91, 71.43 and 118.42, respectively, which all have a p -value that is less than 0.001.

6.3. Features of Correspondence Analysis & Symmetry

Interestingly, for Table 1, closed form expressions exist for many of the features that come from the correspondence analysis of contingency table when examining departures from symmetry. To show this, suppose consider again the matrix of the *Bowker residuals*, \mathbf{S} . For Table 1, this matrix is

$$\mathbf{S} = \begin{pmatrix} 0 & -\frac{C}{\sqrt{2n(40+C)}} & 0 & 0 \\ \frac{C}{\sqrt{2n(40+C)}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The structure of this 4×4 matrix is identical to a 2×2 matrix when removing the vectors of zeros in the last two rows and columns of \mathbf{S} . In Appendix A we show that when \mathbf{S} is of size 2×2 , then applying a SVD to it yields the singular values

$$\lambda_1 = \lambda_2 = \frac{C}{\sqrt{2n(40+C)}} \quad (= s_{21} = -s_{12}) \quad (22)$$

thereby producing two non-trivial, and identical, singular values, and one zero singular value. Thus, the sum-of-squares of these singular values gives, for Table 1, a total inertia of

$$\phi_S^2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \frac{C^2}{n(40 + C)}.$$

Note that this is equivalent to (21) divided by the sample size n . Since $n = 680 + C$ then the total inertia can be expressed in terms of C by

$$\phi_S^2(C) = \frac{C^2}{(680 + C)(40 + C)}$$

so that, for Table 1, $0 \leq \phi_S^2(C) < 1$ for $C \geq 0$.

Since there are only two, equivalent, non-trivial singular values, a two-dimensional correspondence plot will produce an optimal display of the departure from symmetry of Table 1 regardless of the value of C ; the principal inertia of these two dimensions is λ_1 each. Thus, both dimensions of the optimal display will visually describe exactly half of the departure from symmetry which is quantified by $\phi_S^2(C)$. Note that from (22) the first two singular values of \mathbf{S} from Table 1 can be expressed as

$$\lambda_1 = \lambda_2 = \frac{C}{\sqrt{2(680 + C)(40 + C)}}$$

so that

$$\mathbf{D}_\lambda = \frac{C}{\sqrt{2(680 + C)(40 + C)}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

while, the first two non-trivial left and right singular vectors of \mathbf{S} are

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \mathbf{A}\mathbf{J}_2^T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

respectively. The matrix $\hat{\mathbf{D}}$ is

$$\hat{\mathbf{D}} = \begin{pmatrix} \frac{200+C}{2(680+C)} & 0 & 0 & 0 \\ 0 & \frac{400+C}{2(680+C)} & 0 & 0 \\ 0 & 0 & \frac{150}{680+C} & 0 \\ 0 & 0 & 0 & \frac{230}{680+C} \end{pmatrix}$$

so that

$$\hat{\mathbf{D}}^{-1/2} = \sqrt{680 + C} \begin{pmatrix} \sqrt{\frac{2}{200+C}} & 0 & 0 & 0 \\ 0 & \sqrt{\frac{2}{400+C}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{150}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{230}} \end{pmatrix}.$$

Therefore, after some matrix manipulation, the matrix containing the row principal coordinates in the optimal two-dimensional correspondence plot can be expressed in terms of C and some constants (dependent on the cell frequencies of Table 1) such that

$$\mathbf{F} = \frac{C}{\sqrt{40 + C}} \begin{pmatrix} 0 & -\sqrt{\frac{1}{200+C}} \\ -\sqrt{\frac{1}{400+C}} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (23)$$

By following a similar derivation, the matrix of column principal coordinates is

$$\mathbf{G} = \frac{C}{\sqrt{40+C}} \begin{pmatrix} -\sqrt{\frac{1}{200+C}} & 0 \\ 0 & \sqrt{\frac{1}{400+C}} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (24)$$

Thus, (23) and (24) satisfy (16) and (17). The zeros in the third and fourth rows of \mathbf{F} and \mathbf{G} are because there is perfect symmetry between the third and fourth rows and columns of Table 1. So, the position of these points in the correspondence plot lie at the origin, irrespective of the choice of $C \geq 0$. When $C = 0$, all principal coordinates for the row and column categories will lie at the origin since Table 1 exhibits perfect symmetry. Therefore, changes in C will lead to changes in the configuration depicted in the correspondence plot. For example, suppose $C = 50$ so that, for Table 1, $n_{21} = 70$. Then, the key features of the correspondence analysis when an examination of the departures from symmetry are being made using Bowker's statistic can be simply calculated as follows:

$$\begin{aligned} X_S^2 &= \frac{50^2}{40+50} = 27.778 \\ \lambda_1 &= \lambda_2 = \frac{50}{\sqrt{2(680+50)(40+50)}} = 0.138 \\ f_{12} &= -\frac{50}{\sqrt{(40+50) \times (200+50)}} = -0.333 \\ f_{21} &= -\frac{50}{\sqrt{(40+50)(400+50)}} = -0.248 \\ g_{11} &= -\frac{50}{\sqrt{(40+50)(200+50)}} = -0.333 \\ g_{22} &= \frac{50}{\sqrt{(40+50) \times (400+50)}} = 0.248. \end{aligned}$$

A summary of these quantities for $C = 50, 75, 100$ and 150 is given in Table 2.

Table 2. Select output from the correspondence analysis of Table 1 when studying departures from symmetry; $C = 50, 75, 100$ and 150 .

Output	C			
	50	75	100	150
X_S^2	27.778	48.913	71.429	118.421
$\phi_S(C)$	0.038	0.065	0.092	0.143
λ_1	0.138	0.180	0.214	0.267
f_{12}	-0.333	-0.422	-0.488	-0.582
f_{21}	-0.248	-0.321	-0.378	-0.464
g_{11}	-0.333	-0.422	-0.488	-0.582
g_{22}	0.248	0.321	0.378	0.464

Supplementing the numerical summaries that appear in Table 2 are the two-dimensional correspondence plots of Figures 1–4. These figures, and their accompanying numerical features, can be obtained using the R function `bowker.ca.exe()` that is described in Appendix C of this paper. Figure 1 shows such a plot for $C = 50$, Figure 2 is the correspondence plot for $C = 75$, Figure 3 is the plot for $C = 100$ while the two-dimensional correspondence plot of Table 1 when $C = 150$ is given by Figure 4. Based on the statements we made above, it should be of no surprise to see that these four plots show that its two dimensions visually describes exactly half of any departures from symmetry that exist between the variables of

Table 1. Thus, each of the four correspondence plots is optimal in its visual depiction of such departures; since λ_1 and λ_2 are equivalent for all values of C . These correspondence plots also show that C3, C4, R3 and R4 all share the same position at the origin since there is no departure from perfect symmetry for these categories. However, the position of R1, R2, C1 and C2 lie further from the origin as C increases since these categories increasingly deviate from perfect symmetry as C increases.

While R1 and C2 are unaffected by changes in C , since symmetry is assessed by considering the difference between the (2, 1)th and (1, 2)th cell frequencies (or proportions), any departure from symmetry does impact their relative position from each other in the correspondence plot. This can be observed by noting that, along the second dimension,

$$\frac{f_{12}}{g_{22}} = -\sqrt{\frac{400 + C}{200 + C}}$$

is not independent of C . For $C > 0$, f_{12} and g_{22} will always remain on opposite sides of the first dimension with f_{12} being at most $\sqrt{2}$ units further away from the origin than g_{22} ; as $C \rightarrow \infty f_{12} \rightarrow -g_{22}$.

We can gain an understanding of how the position of R1 and C1, say, in the optimal (two-dimensional) correspondence plot compare for $C > 0$. Since $f_{12} < 0$ and $g_{11} < 0$, R1 will lie along the second dimension, while C1 will lie along the first dimension showing that their relative proximity from each other increases as C increases; note that we are not interpreting this row/column proximity in terms of a quantifiable distance measure. However, we can quantify how the relative position of R1 and C1 change by noting the ratio between these two coordinates is $f_{12}/g_{11} = 1$ for all values of C . Therefore, R1 and C1 will move the same number of units away from the origin as C increases.

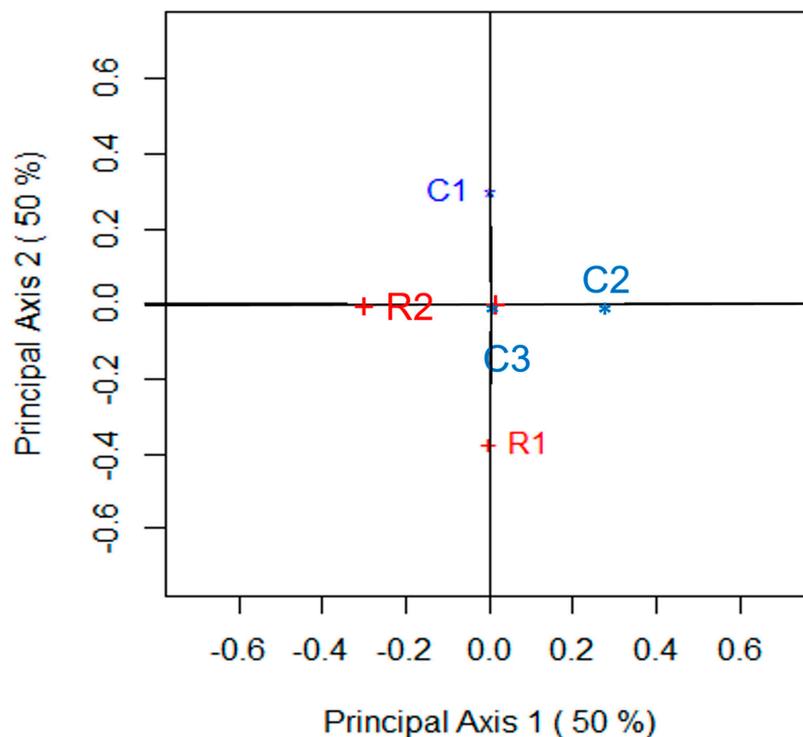


Figure 1. Correspondence plot that visually examines the departure from symmetry for Table 1; $C = 50$.

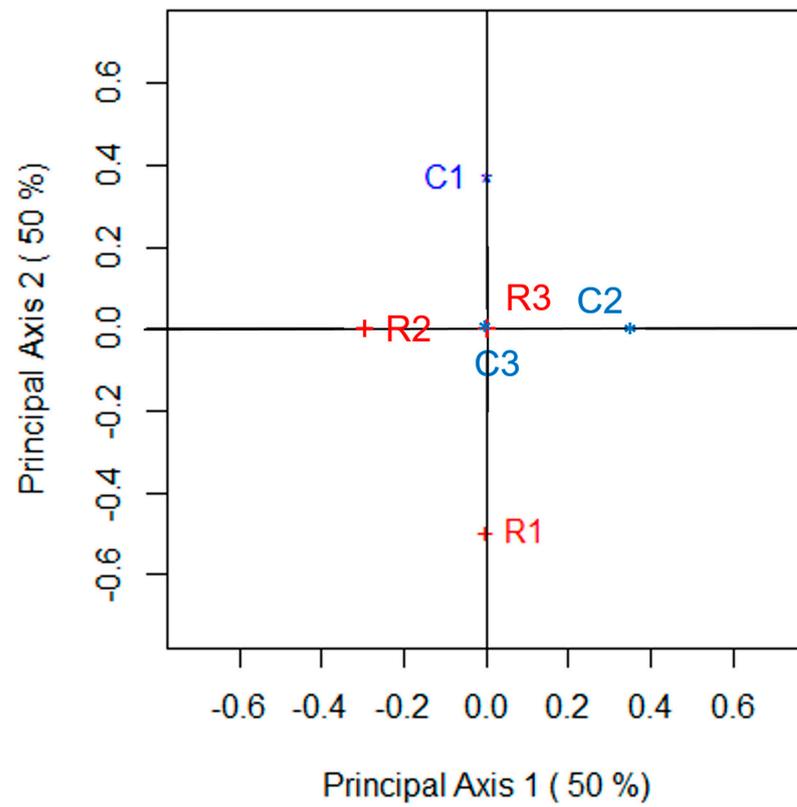


Figure 2. Correspondence plot that visually examines the departure from symmetry for Table 1; $C = 75$.

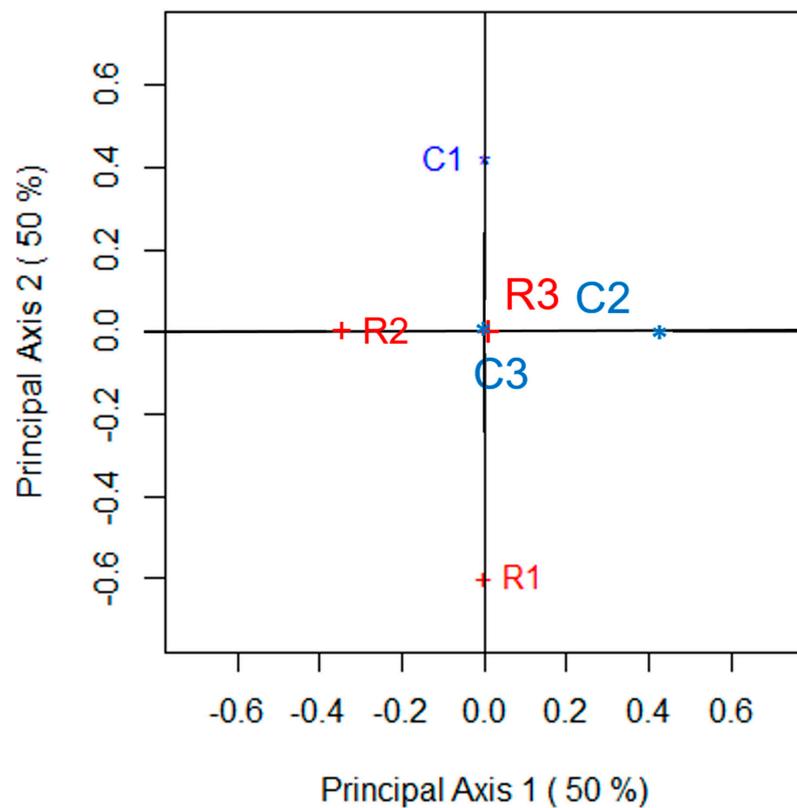


Figure 3. Correspondence plot that visually examines the departure from symmetry for Table 1; $C = 100$.

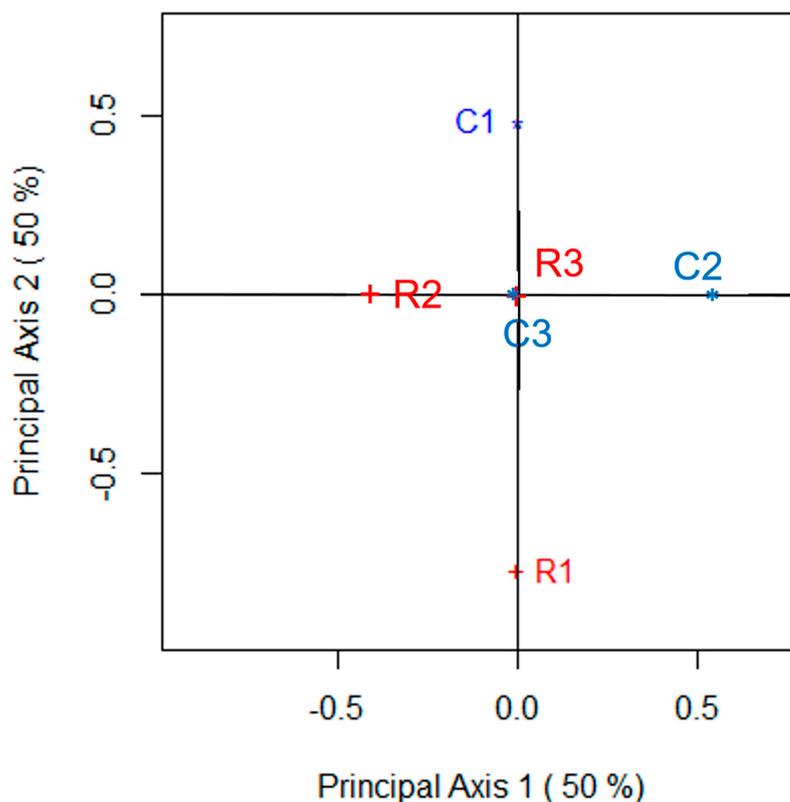


Figure 4. Correspondence plot that visually examines the departure from symmetry for Table 1; C = 150.

7. Example 2 on the Purchase of Decaffeinated Coffee

We now focus our attention on a 5×5 contingency table considered by Agresti [26] (Table 8.5) whose original data came from Grover and Srinivasan [15]; see Table 3. For these data, 541 individuals were surveyed about their choice of purchase of five brands of decaffeinated coffee. Each of the participants were asked to record the brand they bought on their first and subsequent purchase. If every participant of the study bought the same brand of coffee on their first and second purchase then the contingency table would exhibit perfect symmetry. However, this was not the case, and so one may investigate where the departures from symmetry lie. In doing so, one can identify brands that had a similar purchasing pattern on the first and second purchase and those that did not.

Table 3. A simple 5×5 table where we test for symmetry using CA.

First Purchase	Second Purchase					Total
	High Pt (hp)	Taster’s (ta)	Sanka (sa)	Nescafé (ne)	Brim (br)	
High Point (HP)	93	17	44	7	10	171
Taster’s Choice (TC)	9	46	11	0	9	75
Sanka (SA)	17	11	155	9	12	204
Nescafé	6	4	9	15	2	36
Brim	10	4	12	2	27	55
Total	135	82	231	33	60	541

A test of symmetry can be performed and doing so yields a Bowker’s statistic of 20.265. With $5 \times (5 - 1)/2 = 10$ degrees of freedom, this statistic has a *p*-value of 0.027 showing

that there are departures from symmetry present in the data. An evaluation of where departures from symmetry lie in Table 3 can be made by observing the skew-symmetric matrix \mathbf{S} which is

$$\mathbf{S} = \begin{pmatrix} 0 & 0.048 & 0.105 & 0.008 & 0.000 \\ -0.048 & 0 & 0 & -0.061 & 0.042 \\ -0.105 & 0 & 0 & 0 & 0 \\ -0.008 & 0.061 & 0 & 0 & 0 \\ 0 & -0.042 & 0 & 0 & 0 \end{pmatrix}.$$

Observing the relative size of the s_{ij} values of this matrix shows that the greatest source of departure from symmetry appears to be for the coffee brand “High Point” while “Brim” is the coffee brand that deviates least from symmetry (although not perfectly). A visual depiction of the departures from symmetry that are present in Table 3 can be made by performing the correspondence analysis approach described above. Appendix C shows how the R function `bowkerca.exe()` can be used to perform this analysis on Table 3. This analysis gives the following two pairs of non-trivial singular values

$$\lambda_1 = \lambda_2 = 0.1215 \quad \text{and} \quad \lambda_3 = \lambda_4 = 0.0641$$

so that their sum-of-squares gives the total inertia

$$\frac{X_S^2}{n} = 0.1215^2 + 0.1215^2 + 0.0641^2 + 0.0641^2 = 0.0377.$$

A visual depiction of the departure from symmetry is given by Figure 5. The quality of this two-dimensional correspondence is excellent and accounts for

$$100 \times \frac{0.1215^2 + 0.1215^2}{0.0377} = 78.242\%$$

of the departure from symmetry that exists in Table 3. The row and column principal coordinates depicted in Figure 5 are

$$\mathbf{F} = \begin{pmatrix} 0 & -0.213 \\ 0.186 & 0.016 \\ 0.155 & 0 \\ 0.044 & -0.140 \\ -0.007 & 0.075 \end{pmatrix} \quad \mathbf{G} = \begin{pmatrix} -0.213 & 0 \\ 0.016 & -0.186 \\ 0 & -0.155 \\ -0.140 & -0.044 \\ 0.075 & 0.007 \end{pmatrix},$$

respectively, and satisfy (16) and (17).

The following points can be made from the configuration in Figure 5 on departures from symmetry in Table 3. Keeping in mind that this correspondence plot captures slightly more than three-quarters of the departures from symmetry in Table 3

- the purchase of the five coffee brands is different across the first and second purchases and so reflects the departure from symmetry that Bowker’s statistic shows,
- the greatest departure from symmetry is for the coffee brand “High Point” since “HP” and “hp” lie furthest from the origin than any of the four remaining brands. Thus, it is this brand that has undergone the greatest difference in purchasing preference over the two time periods,
- the coffee brands are ordered as follows based on the greatest to least departure from perfect symmetry: “High Point”, “Taster’s Choice”, “Sanka”, “Nescafé” and “Brim”,
- therefore, “Brim” is the coffee brand that has the most similar purchasing pattern across the two time periods when the brands were purchased.

Furthermore, Figure 5 shows that

- the purchasing preferences of the brands “Sanka” and “Taster’s Choice” are very similar on their first purchase as well as on their second purchase. This can be seen

because of the close proximity of “sa” and “tc” on the left of the plot, and “SA” and “TC” on the right of the plot,

- the purchasing preferences of the brands “High Point” and “Nescafé” are similar (although not as similar as “SA” and “TC”) within each of the two purchases.

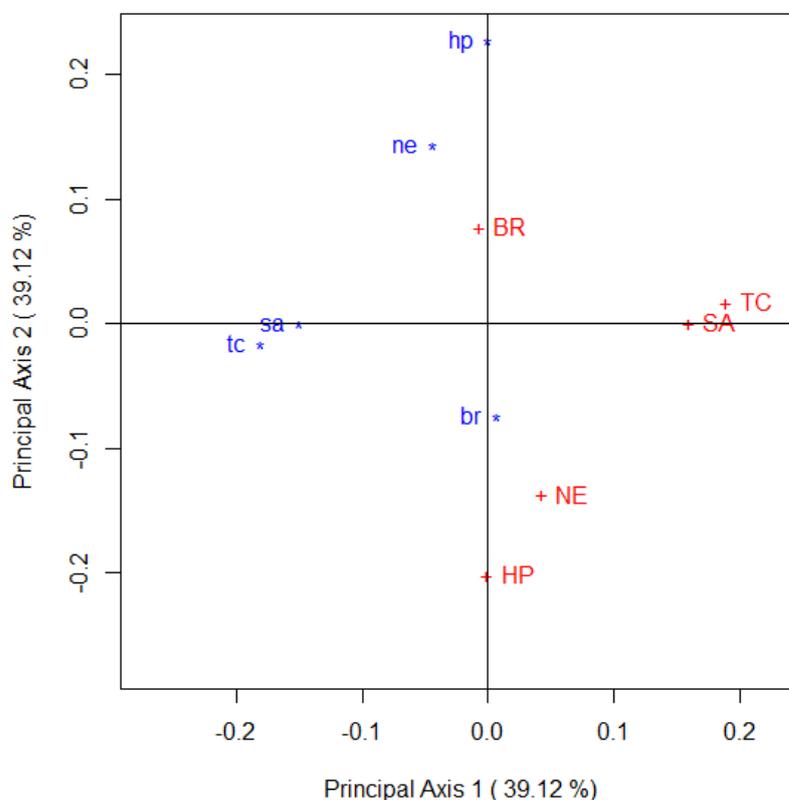


Figure 5. Correspondence plot that visually examines the departure from symmetry for Table 1; $C = 150$.

8. Discussion

When studying departures from symmetry between the categorical variables of a two-way contingency table there are many different techniques that can be considered. A common thread amongst many of them (especially over the past few decades) has been to partition the contingency table into a symmetric part (\mathbf{Y}) and an asymmetric, or skew-symmetric, part (\mathbf{K}) as (1) shows. While such a partition has appeared in the correspondence analysis literature, to the best of our knowledge the above technique is the first to formally link a meaningful measure of asymmetry when visualising departures from symmetry. Specifically, this paper has shown how Bowker’s statistic [14] plays a pivotal role in quantifying such departures in the context of correspondence analysis. Importantly, we also showed that by using Bowker’s statistic, we are able to capture departures from perfect symmetry *relative* to the amount of symmetry that lies between the variables.

In preparing this paper, we considered metrics that differ to those we described above. Like Greenacre [13], we adopted a metric involving the mean row-column marginal proportion. However, since Bowker’s statistic is independent of the row and column marginal information, consideration was given to $\hat{\mathbf{P}} = (\mathbf{P} + \mathbf{P}^T)/2$. While using such a metric does not lead to the exact total inertia, it does provide an excellent approximation to it in some cases. It also provides additional features not available with the metric adopted above and so this is an interesting avenue to pursue in the future.

There are further extensions of the technique described above that can be considered at a later time. One such extension, and one that we raised at the end of Section 3, is to investigate the role of the Cressie–Read family of divergence statistics [21] for visualising

departures from perfect symmetry using correspondence analysis. Doing so then will mean that one can consider “symmetry” (as opposed to “independence”) versions of the special cases of this family of divergence statistics, such as G_S^2 , the Freeman–Tukey statistic [27] and other association measures, as alternatives to Bowker’s X^2 statistic. One can then consider measures of accuracy such as those described by Hubert and Arabie [28] for assessing how different members of this family compare.

Another possible avenue for future research is to extend the above technique in the case where one categorical variable is defined as a predictor variable and the other is its response variable. Such an approach provides a visual means of identifying departures from perfect symmetry using non-symmetrical correspondence analysis. While the approach described above is confined to examining departures from perfect symmetry between two cross-classified categorical variables, another natural extension is to consider adapting the above technique to analyse multi-way contingency tables. There is scope to investigate how this can be achieved in the context of multiple and multi-way correspondence analysis [6]. However, this extension, and the others we described, will be left for consideration at a later date.

Author Contributions: Conceptualization: E.J.B.; Methodology: E.J.B. and R.L.; Software: E.J.B. and R.L.; Artificial and Applied Data Analysis: E.J.B. and R.L.; Writing—original draft preparation: E.J.B.; Writing—review and editing: E.J.B. and R.L.; Visualization: E.J.B. and R.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data in Table 1 is artificial and unique to this paper. The data in Table 3 is from Grover and Srinivasan [15] and also appears in Agresti [26] (Table 8.5).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviation

The following abbreviation is used in this manuscript:

SVD Singular value decomposition

Appendix A. The Singular Values of a 2×2 S Matrix

Suppose we have the following generic 2×2 skew-symmetric matrix

$$\mathbf{S} = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$$

Since $S = 2$, there will be two singular-values for us to determine. To derive them we shall consider the eigen-decomposition of $\mathbf{S}^T \mathbf{S}$ by noting that

$$\begin{aligned} \mathbf{S}^T \mathbf{S} &= \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}^T \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^2 & 0 \\ 0 & a^2 \end{pmatrix}. \end{aligned}$$

Since $\mathbf{S}^T \mathbf{S}$ is a diagonal matrix with identical diagonal elements then it has two identical eigen-values which are

$$\lambda_1^2 = a^2 \quad \text{and} \quad \lambda_2^2 = a^2.$$

In the context of the matrix of the Bowker residuals,

$$a \equiv \frac{1}{\sqrt{2}} \frac{p_{21} - p_{12}}{\sqrt{p_{21} + p_{12}}}.$$

Therefore,

$$\lambda_1^2 = \lambda_2^2 = \frac{1}{2} \frac{(p_{21} - p_{12})^2}{p_{21} + p_{12}}$$

and are two non-trivial, and identical, eigen-vectors of $\mathbf{S}^T \mathbf{S}$. Thus, the two non-trivial singular values of \mathbf{S} are

$$\lambda_1 = \lambda_2 = \frac{1}{\sqrt{2}} \frac{p_{21} - p_{12}}{\sqrt{p_{21} + p_{12}}}$$

so that the total inertia of the 2×2 matrix containing the Bowker residuals is

$$\phi^2 = \lambda_1^2 + \lambda_2^2 = \frac{(p_{21} - p_{12})^2}{p_{21} + p_{12}}$$

and is equivalent to McNemar [17] statistic divided by the sample size. This shows that when performing a correspondence analysis for assessing the departure from symmetry of a 2×2 contingency table, each dimension contributes equally, and to exactly half, of the total inertia.

Appendix B. The Singular Values of a 3×3 \mathbf{S} Matrix

Suppose we now have the generic 3×3 skew-symmetric matrix

$$\mathbf{S} = \begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix}$$

Since $S = 3$, this matrix has exactly two positive singular values and one zero singular value. Here we derive these values in terms of the elements of \mathbf{S} by considering the eigen-decomposition of $\mathbf{S}^T \mathbf{S}$. In doing so

$$\begin{aligned} \mathbf{S}^T \mathbf{S} &= \begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix}^T \begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix} \\ &= \begin{pmatrix} a^2 + b^2 & bc & -ac \\ bc & a^2 + c^2 & ab \\ -ac & ab & b^2 + c^2 \end{pmatrix}. \end{aligned}$$

The eigen-values of $\mathbf{S}^T \mathbf{S}$ in this case are determined by solving the characteristic equation

$$|\mathbf{S}^T \mathbf{S} - \lambda^2 \mathbf{I}_3| = 0$$

where \mathbf{I}_3 is a 3×3 identity matrix. Thus

$$\begin{aligned}
|\mathbf{S}^T \mathbf{S} - \lambda^2 \mathbf{I}_3| &= \begin{vmatrix} a^2 + b^2 - \lambda^2 & bc & -ac \\ bc & a^2 + c^2 - \lambda^2 & ab \\ -ac & ab & b^2 + c^2 - \lambda^2 \end{vmatrix} \\
&= (a^2 + b^2 - \lambda^2) \begin{vmatrix} a^2 + c^2 - \lambda^2 & ab \\ ab & b^2 + c^2 - \lambda^2 \end{vmatrix} \\
&\quad - bc \begin{vmatrix} bc & ab \\ -ac & b^2 + c^2 - \lambda^2 \end{vmatrix} - ac \begin{vmatrix} bc & a^2 + c^2 - \lambda^2 \\ -ac & ab \end{vmatrix} \\
&= (a^2 + b^2 - \lambda^2) [(a^2 + c^2 - \lambda^2)(b^2 + c^2 - \lambda^2) - a^2 b^2] \\
&\quad - bc [bc(b^2 + c^2 - \lambda^2) + a^2 bc] - ac [ab^2 c + ac(a^2 + c^2 - \lambda^2)] \\
&= (a^2 + b^2 - \lambda^2) [(a^2 + c^2)(b^2 + c^2) - \lambda^2(a^2 + c^2) - \lambda^2(b^2 + c^2) \\
&\quad + \lambda^4 - a^2 b^2] - bc [b^3 c + bc^3 - \lambda^2 bc + a^2 bc] \\
&\quad - ac [ab^2 c + a^3 c + ac^3 - \lambda^2 ac] \\
&= (a^2 + b^2 - \lambda^2) [\lambda^4 - \lambda^2(a^2 + b^2 + 2c^2) + (a^2 c^2 + b^2 c^2 + c^4)] \\
&\quad - bc [-\lambda^2 bc + b^3 c + bc^3 + a^2 bc] \\
&\quad - ac [-\lambda^2 ac + ab^2 c + a^3 c + ac^3] \\
&= \lambda^4(a^2 + b^2) - \lambda^2(a^2 + b^2)(a^2 + b^2 + 2c^2) + (a^2 + b^2)(a^2 c^2 + b^2 c^2 + c^4) \\
&\quad - \lambda^6 + \lambda^4(a^2 + b^2 + 2c^2) - \lambda^2(a^2 c^2 + b^2 c^2 + c^4) \\
&\quad + (\lambda^2 b^2 c^2 - b^4 c^2 - b^2 c^4 - a^2 b^2 c^2) \\
&\quad + (\lambda^2 a^2 c^2 - a^2 b^2 c^2 - a^4 c^2 - a^2 c^4) \\
&= -\lambda^6 + \lambda^4(2a^2 + 2b^2 + 2c^2) \\
&\quad - \lambda^2((a^2 + b^2)(a^2 + b^2 + 2c^2) + (a^2 c^2 + b^2 c^2 + c^4) - b^2 c^2 - a^2 c^2) \\
&\quad + [(a^2 + b^2)(a^2 c^2 + b^2 c^2 + c^4) - b^4 c^2 - b^2 c^4 - a^2 b^2 c^2 - a^2 b^2 c^2 \\
&\quad - a^4 c^2 - a^2 c^4] \\
&= -\lambda^6 + \lambda^4(2a^2 + 2b^2 + 2c^2) \\
&\quad - \lambda^2(a^4 + a^2 b^2 + 2a^2 c^2 + a^2 b^2 + b^4 + 2b^2 c^2 + a^2 c^2 + b^2 c^2 + c^4 \\
&\quad - b^2 c^2 - a^2 c^2) + 0.
\end{aligned}$$

Therefore, setting this sextic equation of λ to zero gives

$$\lambda^2 [\lambda^4 - 2\lambda^2(a^2 + b^2 + c^2) + (a^2 + b^2 + c^2 + 2a^2 b^2 + 2a^2 c^2 + 2b^2 c^2)] = 0$$

which is a perfect square so that

$$\lambda^2 [\lambda^4 - 2\lambda^2(a^2 + b^2 + c^2) + (a^2 + b^2 + c^2)^2] = 0$$

or

$$\lambda^2 [\lambda^2 - (a^2 + b^2 + c^2)]^2 = 0.$$

Therefore, there are two positive eigen-values and one zero eigen-value of $\mathbf{S}^T \mathbf{S}$ (when $S = 3$) and they are

$$\lambda_1^2 = \lambda_2^2 = a^2 + b^2 + c^2 \quad \text{and} \quad \lambda_3^2 = 0.$$

This then confirms that the two largest eigen-values of $\mathbf{S}^T \mathbf{S}$, and hence singular values of \mathbf{S} , are identical with a zero third value. In the context of \mathbf{S} ,

$$a \equiv \frac{1}{\sqrt{2}} \frac{p_{21} - p_{12}}{\sqrt{p_{21} + p_{12}}}, \quad b \equiv \frac{1}{\sqrt{2}} \frac{p_{31} - p_{13}}{\sqrt{p_{31} + p_{13}}} \quad \text{and} \quad c \equiv \frac{1}{\sqrt{2}} \frac{p_{32} - p_{23}}{\sqrt{p_{32} + p_{23}}}$$

so that the principal inertia values associated with the dimensions of the optimal correspondence plot are

$$\lambda_1^2 = \lambda_2^2 = \frac{1}{2} \left(\frac{(p_{21} - p_{12})^2}{p_{21} + p_{12}} + \frac{(p_{31} - p_{13})^2}{p_{31} + p_{13}} + \frac{(p_{32} - p_{23})^2}{p_{32} + p_{23}} \right) = \frac{X_s^2}{2n}$$

and

$$\lambda_3^2 = 0.$$

Thus for a 3×3 contingency table, the optimal correspondence plot will consist of two dimensions and each will account for exactly 50% of the total inertia. Note then that it is not surprising that the sum of the three principal inertia values gives the total inertia since

$$\phi^2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \frac{X_s^2}{2n} + \frac{X_s^2}{2n} + 0 = \frac{X_s^2}{n}.$$

Appendix C. R Code

This appendix contains the R function `bowkerca.exe()` that performs a correspondence analysis on an $S \times S$ contingency table where the departure from perfect symmetry is assessed using Bowker's X^2 statistic—see (6). The arguments of the function are

- `N`—the two-way contingency table of size $S \times S$, where $S > 2$,
- `scaleplot`—rescales the limit of the axes used to construct the two-dimensional correspondence plot. By default, `scaleplot = 1.2`,
- `dim1`—the first dimension of the correspondence plot. By default, `dim1 = 1` so that the first dimension is depicted horizontally, and
- `dim2`—the second dimension of the correspondence plot. By default, `dim2 = 2` so that the second dimension is depicted vertically

```
bowkerca.exe <- function(N, scaleplot = 1.2, dim1 = 1, dim2 = 2){
  S <- nrow(N)           # Number of rows & columns of the table
  Inames <- dimnames(N)[1] # Row category names
  Jnames <- dimnames(N)[2] # Column category names

  n <- sum(N)           # Total number of classifications in the table
  p <- N * (1/n)       # Matrix of joint relative proportions

  pidot <- apply(p, 1, sum) # Row marginal proportions
  pdotj <- apply(p, 2, sum) # Column marginal proportions

  dI <- diag(pidot, nrow = S, ncol = S)
  dJ <- diag(pdotj, nrow = S, ncol = S)
  dIJ <- 0.5*(dI + dJ)

  # Constructing the matrix of Bowker residuals

  s <- matrix(0, nrow = S, ncol = S)

  for (i in 1:S){
    for (j in 1:S){
      s[i,j] <- (p[i,j] - (p[i,j] + p[j,i])/2) / sqrt((p[i,j] + p[j,i])/2)
    }
  }
}
```

```

}

dimnames(s) <- list(paste(Inames[[1]]), paste(Jnames[[1]])

# Applying a singular value decomposition (SVD) to the matrix of
# Bowker residuals

sva <- svd(s)

d <- sva$d
dmu <- diag(sva$d)

#####
#
# Principal Coordinates
#
#####

# Row principal coordinates

f <- solve(dIJ^0.5) %*% sva$u %*% dmu
dimnames(f) <- list(paste(Inames[[1]]), paste(1:S))

# Column principal coordinates

g <- solve(dIJ^0.5) %*% sva$v %*% dmu
dimnames(g) <- list(paste(Jnames[[1]]), paste(1:S))

#####
#
# Calculating the total inertia, Bowker's chi-squared
# statistic, its p-value and the percentage contribution
# of the axes to the inertia
#
#####

Principal.Inertia <- diag(t(f) %*% dIJ %*% f)
Total.Inertia <- sum(Principal.Inertia)
Bowker.X2 <- n * Total.Inertia # Bowker's Chi-squared statistic
Perc.Inertia <- (Principal.Inertia/Total.Inertia) * 100
Cumm.Inertia <- cumsum(Perc.Inertia)
Inertia <- cbind(Principal.Inertia, Perc.Inertia, Cumm.Inertia)
dimnames(Inertia)[1] <- list(paste("Axis", 1:S, sep = " "))
p.value <- 1 - pchisq(Bowker.X2, S * (S - 1)/2)

#####
#
# Here we construct the 2-D correspondence plot
#
#####

par(pty = "s")

```

```

plot(0, 0, pch = " ",
     xlim = scaleplot*range(f[, dim1], f[, dim2], g[, dim1], g[, dim2]),
     ylim = scaleplot*range(f[, dim1], f[, dim2], g[, dim1], g[, dim2]),
     xlab = paste("Principal Axis", dim1, "(", round(Perc.Inertia[dim1],
           digits = 2), "%)",
     ylab = paste("Principal Axis", dim2, "(", round(Perc.Inertia[dim2],
           digits = 2), "%)",
)

points(f[, dim1], f[, dim2], pch = "+", col = "red")
text(f[, dim1], f[, dim2], labels = Inames[[1]], pos = 4, col = "red")

points(g[, dim1], g[, dim2], pch = "*", col = "blue")
text(g[, dim1], g[, dim2], labels = Jnames[[1]], pos = 2, col = "blue")

abline(h = 0, v = 0)

list(N = N,
     s = round(s, digits = 3),
     f = round(f, digits = 3),
     g = round(g, digits = 3),
     Bowker.X2 = round(Bowker.X2, digits = 3),
     P.Value = round(p.value, digits = 3),
     Total.Inertia = round(Total.Inertia, digits = 3),
     Inertia = round(Inertia, digits = 3)
)
}

```

The numerical summaries that are produced from this function are

- the contingency table under investigation, N ,
- the matrix of Bowker residuals, s , where the elements are defined by (7),
- the matrix of row principal coordinates, f , and column principal coordinates, g , defined by (14) and (15), respectively,
- Bowker's chi-squared statistic defined by (6), $Bowker.X2$, and its p-value, $P.Value$, and
- the principal inertia value for each of the M dimensions, $Principal.Inertia$, the percentage of the total inertia accounted for by each of these dimensions, $Perc.Inertia$, and the cumulative percentage of the M principal inertia values, $Cumm.Inertia$.

Therefore, when `coffee.dat` is the R object assigned to Table 3 so that

```

> coffee.dat <- matrix(c(93, 9, 17, 6, 10, 17, 46, 11, 4, 4, 44, 11, 155,
+ 9, 12, 7, 0, 9, 15, 2, 10, 9, 12, 2, 27), nrow = 5)
> dimnames(coffee) <- list(paste(c("HP", "TC", "SA", "NE", "BR")),
+ paste(c("hp", "tc", "sa", "ne", "br")))
>

```

then the function produces the correspondence plot of Figure 5 and the following numerical summaries

```

> bowkerca.exe(coffee)
$N
  hp tc  sa ne br
HP 93 17 44 7 10
TC  9 46 11 0  9
SA 17 11 155 9 12

```

```
NE 6 4 9 15 2
BR 10 4 12 2 27
```

```
$s
      hp      tc      sa      ne      br
HP  0.000  0.048  0.105  0.008  0.000
TC -0.048  0.000  0.000 -0.061  0.042
SA -0.105  0.000  0.000  0.000  0.000
NE -0.008  0.061  0.000  0.000  0.000
BR  0.000 -0.042  0.000  0.000  0.000
```

```
$f
      1      2      3      4 5
HP  0.000 -0.213  0.000  0.043 0
TC  0.186  0.016 -0.135  0.023 0
SA  0.155  0.000  0.059  0.000 0
NE  0.044 -0.140 -0.020 -0.193 0
BR -0.007  0.075  0.017  0.103 0
```

```
$g
      1      2      3      4 5
hp -0.213  0.000 -0.043  0.000 0
tc  0.016 -0.186 -0.023 -0.135 0
sa  0.000 -0.155  0.000  0.059 0
ne -0.140 -0.044  0.193 -0.020 0
br  0.075  0.007 -0.103  0.017 0
```

```
$Bowker.X2
[1] 20.412
```

```
$P.Value
[1] 0.026
```

```
$Total.Inertia
[1] 0.038
```

```
$Inertia
      Principal.Inertia Perc.Inertia Cumm.Inertia
Axis 1          0.015         39.121         39.121
Axis 2          0.015         39.121         78.242
Axis 3          0.004         10.879         89.121
Axis 4          0.004         10.879        100.000
Axis 5          0.000          0.000        100.000
```

```
>
```

References

1. Agresti, A. *Categorical Data Analysis*, 2nd ed.; Wiley: New York, NY, USA, 2002.
2. Anderson, E.B. *The Statistical Analysis of Categorical Data*; Springer: Berlin/Heidelberg, Germany, 1991.
3. Bove, G. Asymmetrical multidimensional scaling and correspondence analysis for square tables. *Stat. Appl.* **1992**, *4*, 587–598.
4. De Falguerolles, A.; van der Heijden, P.G.M. Reduced rank quasi-symmetry and quasi-skew symmetry: A generalized bi-linear model approach. *Ann. Fac. Sci. Toulouse* **2002**, *11*, 507–524. [[CrossRef](#)]
5. Iki, K.; Yamamoto, K.; Tomizawa, S. Quasi-diagonal exponent symmetry model for square contingency tables with ordered categories. *Stat. Probab. Lett.* **2014**, *92*, 33–38. [[CrossRef](#)]

6. Tomizawa, S. Two kinds of measures of departure from symmetry in square contingency tables having nominal categories. *Stat. Sin.* **1994**, *4*, 325–334.
7. Yamamoto, H. A measure of departure from symmetry for multi-way contingency tables with nominal categories. *Jpn. J. Biom.* **2004**, *25*, 69–88. [[CrossRef](#)]
8. Yamamoto, K.; Shimada, F.; Tomizawa, S. Measure of departure from symmetry for the analysis of collapsed square contingency tables with ordered categories. *J. Appl. Stat.* **2015**, *42*, 866–875. [[CrossRef](#)]
9. Arellano-Valle, R.B.; Contreras-Reyes, J.E.; Stehlik, M. Generalized skew-normal negentropy and its application to fish condition factor time series. *Entropy* **2017**, *19*, 528. [[CrossRef](#)]
10. Nishisato, S. *Optimal Quantification and Symmetry*; Springer: Singapore, 2022.
11. Constantine, A.G.; Gower, J.C. Graphical representation of asymmetry. *Appl. Stat.* **1978**, *27*, 297–304. [[CrossRef](#)]
12. Gower, J.C. The analysis of asymmetry and orthogonality. In *Recent Developments in Statistics*; Barra, J.R., Brodeau, F., Romer, G., van Cutsem, B., Eds.; North-Holland: Amsterdam, The Netherlands, 1977; pp. 109–123.
13. Greenacre, M. Correspondence analysis of square asymmetric matrices. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2000**, *49*, 297–310. [[CrossRef](#)]
14. Bowker, A.H. A test for symmetry in contingency tables. *J. Am. Stat. Assoc.* **1948**, *43*, 572–598. [[CrossRef](#)] [[PubMed](#)]
15. Grover, R.; Srinivasan, V. A simultaneous approach to market segmentation and market structuring. *J. Mark. Res.* **1987**, *24*, 129–153. [[CrossRef](#)]
16. Beh, E.J.; Lombardo, R. *Correspondence Analysis: Theory, Practice and New Strategies*; Wiley: Chichester, UK, 2014.
17. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [[CrossRef](#)] [[PubMed](#)]
18. Bishop, Y.M.; Fienberg, S.E.; Holland, P.W. *Discrete Multivariate Analysis: Theory and Practice*; MIT Press: Cambridge, MA, USA, 1975.
19. Lancaster, H.O. *The Chi-squared Distribution*; Wiley: Sydney, Australia, 1969.
20. Plackett, R.L. *The Analysis of Categorical Data*; Charles Griffin and Company, Limited: London, UK, 1974.
21. Cressie, N.A.C.; Read, T.R.C. Multinomial goodness-of-fit tests. *J. R. Stat. Soc. Ser. B* **1984**, *46*, 440–464. [[CrossRef](#)]
22. Beh, E.J.; Lombardo, R. *Correspondence Analysis and the Cressie–Read Family of Divergence Statistics*; National Institute for Applied Statistics Research Australia (NIASRA) Working Paper Series; 2022. Available online: <https://www.uow.edu.au/niasra/publications/> (accessed on 19 May 2022).
23. Ward, R.C.; Gray, L.J. Eigensystem computation for skew-symmetric matrices and a class of symmetric matrices. *Acm Trans. Math. Softw.* **1978**, *4*, 278–285. [[CrossRef](#)]
24. Murnaghan, F.D.; Wintner, A. A canonical form for real matrices under orthogonal transformations. *Proc. Natl. Acad. Sci. USA* **1931**, *17*, 417–420. [[CrossRef](#)] [[PubMed](#)]
25. Paardekooper, M.H.C. An eigenvalue algorithm for skew-symmetric matrices. *Numer. Math.* **1971**, *17*, 189–202. [[CrossRef](#)]
26. Agresti, A. *An Introduction to Categorical Data Analysis*, 3rd ed.; Wiley: New York, NY, USA, 2019.
27. Freeman, M.F.; Tukey, J.W. Transformations related to the angular and square root. *Ann. Math. Stat.* **1950**, *21*, 607–611. [[CrossRef](#)]
28. Hubert, L.; Arabie, P. Comparing partitions. *J. Classif.* **1985**, *2*, 193–218. [[CrossRef](#)]