MULTIPLE CORRESPONDENCE ANALYSIS: A BROAD OVERVIEW

By Soham Bonnerjee, Sagnik Roy, Soumen Ghosh

Indian Statistical Institute, Kolkata

Guide : S.S.Mukherjee, *Indian Statistical Institute, Kolkata*

In statistics, Multiple Correspondence Analysis (MCA) is a data analysis technique for nominal categorical data, used to detect and represent underlying structures in a data set. It does this by representing data as points in a low-dimensional Euclidean space. In this article we aim to briefly discuss the motivation and the theories behind MCA, and also explore the different realms of Statistics this technique is applicable in.

1 Introduction

Multiple Correspondence Analysis was developed in France by the mathematician Jean-Paul Benzécri[1], and popularized by Bourdieu[2] outside France. It is a generalization of Simple Correspondence Analysis, for the latter is primarily applicable to 2-way contingency tables, whereas MCA tackles the more general problem of relation between categories in higher-dimensional tables. Multiple Correspondence Analysis, also, in essence, gives a "factor" variable for each level of each category, of which we usually take the first two coordinates to help us visualize the categories in a multidimensional table in a two-dimensional biplot. Thus it can be viewed as a counterpart of PCA for Categorical data.

In Section 2, we briefly discuss how Multiple Correspondence Analysis is performed. In Section 3, we express MCA as a generalization of Simple Correspondence Analysis. In section 3, we analyze a real data using MCA to get a hands-on feeling of the algorithm. In section 5, we briefly take up the topic of eigenvalue correction. Finally in Section 6, 7 and 8, we take up three more real life data pertaining to three different statistical problems, and show how MCA helps us in achieving the goal.

2 Multiple Correspondence Analysis Algorithm

Consider a contingency table with n observation and Q categories. The i-th category has k_i levels. Define $J = \sum_{i=1}^{Q} k_i$. MCA can be simply performed by applying the CA algorithm to either a *Super-Indicator matrix* (also called complete disjunctive table – CDT) or a *Burt Matrix* formed from these variables.

2.1 MCA using Super-Indicator Matrix

- For level i, form a $n \times k_i$ indicator matrix X_i as follows:

$$(X_i)_{pq} = \mathbb{I}\,(p\text{-th observation belongs to } q\text{-th level of the category})$$

- Form the $n \times J$ super indicator matrix X as follows:

$$X = (X_1 : X_2 : \cdots : X_Q)$$

- Let $N = \mathbf{1}^T X \mathbf{1} = nQ$ be the sum of all the entries in X. Define $Z = \dfrac{X}{N}$. The row sum of Z is $\mathbf{r^X} = \dfrac{1}{n}\mathbf{1}$. Let the vector of marginals of category i be $\mathbf{c_i}$. Then the vector of column sums of Z is $\mathbf{c^X} = \dfrac{1}{N}\left(\mathbf{c_1} : \mathbf{c_2} : \cdots : \mathbf{c_Q}\right)^T$. Also let $\mathbf{D^X_r} = \text{Diag}(\mathbf{r^X}) = \dfrac{1}{n}\mathbf{I}_{n\times n}$ and $\mathbf{D^X_c} = \text{Diag}(\mathbf{c^X})$
- Perform a Singular Value Decomposition of

$$C = \mathbf{D_r^X}^{-\frac{1}{2}}\left(Z - \mathbf{r^X}(\mathbf{c^X})^T\right)\mathbf{D_c^X}^{-\frac{1}{2}} = \sqrt{n}\left(Z - \mathbf{r^X}(\mathbf{c^X})^T\right)\mathbf{D_c^X}^{-\frac{1}{2}}$$

:

$$C = \Gamma\Lambda\Delta^T$$

.

- The row and respectively column factor scores are obtained by:

$$\mathbf{F} = \mathbf{D_r^X}^{-\frac{1}{2}}\Gamma\Lambda \qquad\qquad \mathbf{G} = \mathbf{D_c^X}^{-\frac{1}{2}}\Delta\Lambda$$

These are called *Principal Coordinates.* Note that the number of factors would be $s = J - Q$. The *Standard Coordinates* are given by:

$$\mathbf{F_s} = \mathbf{D_r^X}^{-\frac{1}{2}}\Gamma \qquad\qquad \mathbf{G_s} = \mathbf{D_c^X}^{-\frac{1}{2}}\Delta$$

- Let the singular values be $\lambda_1, \cdots, \lambda_s$. Then the squared singular values, i.e the eigenvalues of CC^T : $\lambda_1^2, \cdots, \lambda_s^2$ are called Principal Inertia of the matrix X.

2.2 MCA using Burt Matrix

- Consider the Burt Matrix $B = X^T X$. Note that this matrix is symmetric, and has the vector of row sums = vector of column sums = $NQ\mathbf{c}^X$. We apply the similar algorithm as above to get row and column factor scores. Let all the corresponding quantities for the MCA using **B** be denoted by a subscript B.

This has the following properties:

- *Since* B *is symmetric, the solution for the rows and for the columns is identical.*
- *The analysis of* B *only gives a solution for the response categories (that is, what were previously the columns of* **X**.
- *The singular values in the analysis of* **B** *are also eigenvalues.*

Proof: Note that $\mathbf{C}_B = (\mathbf{C}_B)^T$. Hence

$$\mathbf{C}_B = \Gamma_B\Lambda_B\Delta_B^T$$

implies that

$$\mathbf{C}_B^2 = \Gamma_B\Lambda_B^2\Gamma_B^T = \Delta_B\Lambda_B^2\Delta_B^T$$

which proves that $\Gamma_B = \Delta_B$ proving the first two properties. Further we get:

$$\mathbf{C}_B = \Gamma_B\Lambda_B\Gamma_B^T$$

which implies Λ_B is indeed the diagonal matrix containing Eigen Values. ■

- *The standard coordinates of the row (equivalent to columns) of* B, *are identical to the standard coordinates of the columns of* $\mathbf{X}$, *and the principal inertias of* $\mathbf{B}$ *are the squares of those of* $\mathbf{X}$.

  Proof: Let $R = {\mathbf{D_r^X}}^{-1} Z = nZ = \frac{1}{Q}X$, and $S = {\mathbf{D_c^X}}^{-1} Z^T = (nQ\mathbf{D_c^X})^{-1}X^T$.
  Consider the relation between F and G as shown in Härdle-Simar[3]

$$
\begin{aligned}
&F = {\mathbf{D_r^X}}^{-1} ZG\Lambda^{-1} \\
\Longrightarrow\ &F = RG\Lambda^{-1} \\
\Longrightarrow\ &RG = F\Lambda \\
\Longrightarrow\ &RG_s\Lambda = F\lambda \\
\Longrightarrow\ &RG_s = F \\
\Longrightarrow\ &SRG_s = SF = G\Lambda = G_s\Lambda^2
\end{aligned}
$$

  where $SF = G\Lambda$ follows similarly as $RG = F\Lambda$. Note that $\mathbf{D_c^B} = \mathbf{D_c^X}$. Hence,

$$
\begin{aligned}
&SRG_s = G_s\Lambda^2 \\
\Longrightarrow\ &(nQ\mathbf{D_c^X})^{-1}X^T\frac{1}{Q}XG_s = G_s\Lambda^2 \\
\Longrightarrow\ &(nQ^2\mathbf{D_c^X})^{-1}X^TXG_s = G_s\Lambda^2 \\
\Longrightarrow\ &(NQ\mathbf{D_c^B})^{-1}BG_s = G_s\Lambda^2 \\
\Longrightarrow\ &S_BG_s = G_s\Lambda^2
\end{aligned}
$$

  which is exactly the relation between the standard coordinates of B, when $(G_B)_s = G_s$ and $\Lambda_B = \Lambda^2$. Thus the standard coordinates of the analysis of Burt Matrix are same as that of X, and the principal inertias are squared to that of X. ■

3 Relation with CA

Consider a $j_1 \times j_2$ Bi-variate Contingency table $\mathbf{M}$. Form the super-indicator matrix $\mathbf{X} = (\mathbf{X_1} : \mathbf{X_2})$. Note that

$$\mathbf{M} = \mathbf{X_1}^T\mathbf{X_2}$$

Let $\mathbf{1}'\mathbf{M}\mathbf{1} = n$. Let the vector of row marginals of $\mathbf{P} = \frac{1}{n}\mathbf{M}$ be $\mathbf{r}$, and the vector of column marginals be $\mathbf{c}$. Then

$$\mathbf{r^X} = \frac{1}{n}\mathbf{1}$$

and

$$\mathbf{c^X} = \frac{1}{2}\begin{bmatrix}\mathbf{r}\\ \mathbf{c}\end{bmatrix}$$

Then $D_r^X = \frac{1}{n}\mathbb{I}$ and $D_c^X = \frac{1}{2}\mathrm{Diag}(\mathbf{D_r}, \mathbf{D_c})$.

To differentiate between standard coordinates obtained from the MCA of X from that from CA of $\mathbf{M}$, we write F_s, G_s as F_s^X and G_s^X, and the corresponding quantities as $F_s^{\mathbf{M}}$ and $G_s^{\mathbf{M}}$. Similarly we redefine all the corresponding quantities.

Then as in the proof of the last property, we have that:

$$
\begin{aligned}
& S^X R^X G_s^X = G_s^X \Lambda_X^2 \\
\implies & 2 \begin{bmatrix} \mathbf{D_r}^{-1} & 0 \\ 0 & \mathbf{D_c}^{-1} \end{bmatrix} \frac{1}{2n} X^T \frac{1}{2} X G_s^X = G_s^X \Lambda_X^2 \\
\implies & \frac{1}{2n} \begin{bmatrix} \mathbf{D_r}^{-1} & 0 \\ 0 & \mathbf{D_c}^{-1} \end{bmatrix} \begin{bmatrix} X_1^T X_1 & X_1^T X_2 \\ X_2^T X_1 & X_2^T X_2 \end{bmatrix} \begin{bmatrix} \mathbf{G_s^{X_1}} \\ \mathbf{G_s^{X_2}} \end{bmatrix} = \begin{bmatrix} \mathbf{G_s^{X_1}} \\ \mathbf{G_s^{X_2}} \end{bmatrix} \Lambda_X^2 \\
\implies & \frac{1}{2n} \begin{bmatrix} \mathbf{D_r}^{-1} & 0 \\ 0 & \mathbf{D_c}^{-1} \end{bmatrix} \begin{bmatrix} n\mathbf{D_r} & n\mathbf{P} \\ n\mathbf{P}^T & n\mathbf{D_c} \end{bmatrix} \begin{bmatrix} \mathbf{G_s^{X_1}} \\ \mathbf{G_s^{X_2}} \end{bmatrix} = \begin{bmatrix} \mathbf{G_s^{X_1}} \\ \mathbf{G_s^{X_2}} \end{bmatrix} \Lambda_X^2
\end{aligned}
$$

This simplifies to the following two equations:

$$
G_s^{X_1} + \mathbf{D_r}^{-1} \mathbf{P} G_s^{X_2} = 2 G_s^{X_1} \Lambda_X^2 \tag{1}
$$

$$
G_s^{X_2} + \mathbf{D_c}^{-1} \mathbf{P}^T G_s^{X_1} = 2 G_s^{X_2} \Lambda_X^2 \tag{2}
$$

Multiplying (1) by $\mathbf{D_c}^{-1}\mathbf{P}^T$ and using the expression for $\mathbf{D_c}^{-1}\mathbf{P}^T G_s^{X_1}$, we have:

$$
\begin{aligned}
\mathbf{D_c}^{-1} \mathbf{P}^T \mathbf{D_r}^{-1} \mathbf{P} G_s^{X_2} &= G_s^{X_2} (2\Lambda_X^2 - \mathbf{I})(2\Lambda_X^2 - \mathbf{I}) \\
\implies S^{\mathbf{M}} R^{\mathbf{M}} G_s^{X_2} &= G_s^{X_2} (2\Lambda_X^2 - \mathbf{I})^2
\end{aligned}
$$

Similarly, we have,

$$
R^{\mathbf{M}} S^{\mathbf{M}} G_s^{X_1} = G_s^{X_1} (2\Lambda_X^2 - \mathbf{I})^2
$$

Now for CA of $\mathbf{P}$, the standard coordinates would be obtained by solving:

$$
\begin{aligned}
S^{\mathbf{M}} R^{\mathbf{M}} G_s^{\mathbf{M}} &= G_s^{\mathbf{M}} \Lambda_M^2 \\
R^{\mathbf{M}} S^{\mathbf{M}} F_s^{\mathbf{M}} &= F_s^{\mathbf{M}} \Lambda_M^2
\end{aligned}
$$

Hence, we have that $F_s^{\mathbf{M}} = G_s^{X_1}$, and $G_s^{\mathbf{M}} = G_s^{X_2}$, and $\Lambda_M^2 = (2\Lambda_X^2 - \mathbf{I})^2$. Thus:

$$
\lambda_{i\mathbf{M}} = \pm(2\lambda_{iX}^2 - 1) = \pm(2\lambda_{iB} - 1)
$$

4 Application MCA: A real-life example

In this section, we take up a real data, and analyze it as a first hand exploratory analysis using MCA to show how it helps us interpreting the data.

We analyze the `Cars93` data from **MASS** package in R. This dataset is from 93 Cars on Sale in the USA in 1993.

The data has 4 qualitative variables:

- `Type`: Describes the car-type. One of 6 types: Compact, Large, Midsize, Small, Sporty, Van.
- `AirBags`: Indicates the existence of air-bags. One of 3 types: Driver and Passenger, Driver, None.

- `DriveTrain`: Drive Train type. One of 3 factors: 4WD, Front Wheel, Rear Wheel.
- `Origin`: Country of origin. One of two factors: USA, non-USA.

We apply MCA on the super-indicator matrix. Note that there will be $14 - 4 = 10$ factors. The screeplot looks as follows:

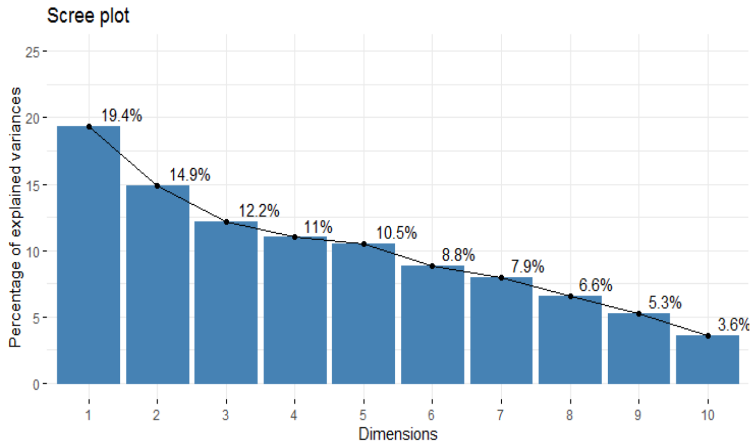


Figure 1: Variances Explained by Dimensions

Although apparently the first two dimensions explain only 34.3% of the inertia, for ease of visualization we take the first two dimensions.Let us have a look at the biplot:

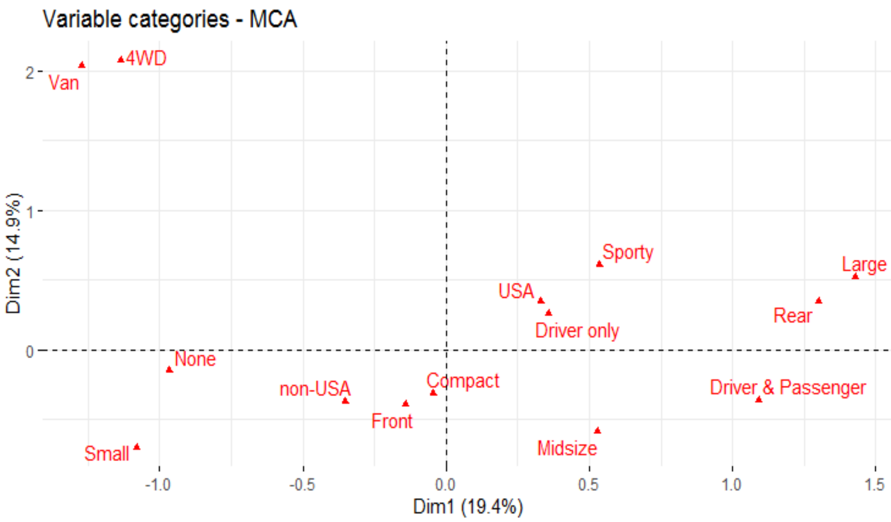


Figure 2: Biplot

From the biplot, we can say the following things about the sample:

- With reference to the principal plane, Car-type Van and Drive-Train 4WD have occurred together.

- Occupying the lower left corner are "Small" cars with No-airbags.
- Cars with airbags for both driver and passenger are mostly midsize in this sample.
- Cars with rear wheel drive-train tend to be large.

5 Correction of Eigenvalues for MCA

MCA codes data by creating several binary columns for each variable with the constraint that one and only one of the columns gets the value 1. This coding scheme creates *artificial* additional dimensions because one categorical variable is coded with several columns. As a consequence, the inertia (i.e the variance) of the solution space is artificially inflated, and therefore the percentage explained by the the first two dimensions is severely *underestimated*.

It can be shown that all the factors with an eigenvalue less than or equal to $\frac{1}{Q}$ simply code these additional dimensions. The correction formula suggested by Greenacre (1993) takes this into account. Specifically, if λ_l^2 denote the eigenvalues (square of the singular values) obtained from the analysis of indicator matrix, then, the corrected eigenvalues, λ_c^2 is given by:

$$\lambda_c^2 = \left[\frac{Q}{Q-1}\left(\lambda_l^2 - \frac{1}{Q}\right)\right]^2 \mathbb{I}\left(\lambda_l^2 > \frac{1}{Q}\right)$$

Traditionally, the percentages of inertia are computed by dividing each eigenvalue by the sum of the eigenvalues, and this approach could be used here also. However, it will give an optimistic estimation of the percentage of inertia. A better estimation of the inertia has been proposed by Greenacre (1993) who suggested in-stead to evaluate the percentage of inertia relative to the average inertia of the off-diagonal blocks of the Burt matrix. This average inertia, I is given by:

$$I = \frac{Q}{Q-1}\left(\sum_l \lambda_l^4 - \frac{J-Q}{Q^2}\right)$$

The percentage of inertia explained by each corrected eigenvalue is given by:

$$\tau_c = \frac{\lambda_c^2}{I}$$

Applying this correction to the above data:

```
car.mca<-MCA(data, ncp=2, graph=FALSE)
e<-car.mca$eig
# average inertia
K<-(4/3)*(sum((e[,1])^2)-10/16)
# corrected eigenvalues
for(i in 1:10){
  if(e[i,1]>1/4){
    e[i,1]<-((4/3)*(e[i,1]-1/4))^2
  }
  else {e[i,1]<-0}
}
# inertia explained by first two eigenvalues
e[1,1]/K
e[2,1]/K
```

We see that about 58.85% of inertia was explained by first dimension and 15% of inertia is explained by the second dimension, implying that the inferences drawn from biplot was not that off. Next we see an application of MCA in a slightly different paradigm.

6 Application of MCA in Conjoint Analysis

'Conjoint analysis' is a survey-based statistical technique used mainly in market research that helps determine how people value different attributes (feature, function, benefits) that make up an individual product or service. It is also really prevalent is psychological studies.

In this section, we take up the `immigrationconjoint` data from **cjoint** package. 1396 Americans were presented with five questions to gauge their outlook to immigration. Each question described 5 features (sex, reason for application, job experience, job plan, language skills) each of two possible immigrants, and the respondents has to choose one of them for allowed immigration. The 5 features are:

- Sex: Male or Female
- Reason of Applicaion: **Reunite** with family, seek better **Job**, Escape **Persecution**.
- Job Experience: None ("Exp0"), 1-2 Yrs ("Exp2"), 3-5 Yrs ("Exp5"), 5+ Yrs ("Exp10").
- Job Plans: None ("PlanN"), Will Look for Work ("PlanW"), Contract with Employer ("PlanC"), Interview with employer ("PlanI").
- Language Skills: Fluent English ("LF"), Broken English ("LB"), Tried English but unable ("LU"), Used Interpreter ("LI").

Each respondent makes 5 choices; hence, for each respondent, one can calculate how many times any particular variable is chosen. These totals (out of 5) are marginals of the incomplete contingency table. They are analogous to marginal means or part-worths in traditional conjoint analysis. There are 17 column totals that represent the attributes in the input matrix. To ensure preliminary comparability among all variables in the analysis, we divided the totals by 5 to obtain fuzzy- coded pairs of attribute levels. We denote the resulting by 1396×17 pseudo-indicator matrix as $\mathbf{X}$. Obviously, all row totals of $\mathbf{X}$ are equal to 5.

6.1 MCA on Pseudo-indicator Matrix

To interpret an MCA (CA) solution, one calculates the so-called absolute contributions and squared correlations for each axis. But for a fuzzy-coded indicaor matrix, the total inertia $\mathrm{In}(j)$ is no longer equal to $(J/p)-1$. But, as Greenacre(1987)[4] suggested, we can still interpret the biplot and obtain meaningful relations about the variables.

The general principle of fuzzy coding schemes is that a constant unit is spread across five categories of the variable, possibly in negative amounts as in the last example. The geometry is a simple generalization of that for indicator matrices and usually amounts to talking about weighted averages of category points instead of simple averages. Hence the analysis can be thought of as a multiple correspondence analysis in which the category points are constrained to be regularly ordered and spaced on a straight line. Thus, the biplot can be interpreted similarly as in the case of super-indicator matrices.

Let's have a look at the biplot.

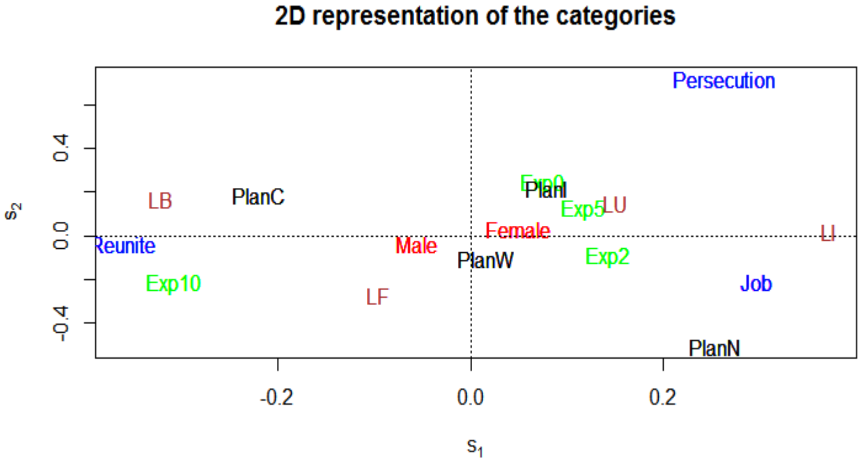


Figure 3: Biplot

From this biplot, evidently respondents who found immigration for reuniting with family to be acceptable, also tended to choose immigrants with broken English, 5+ years of experience and contract with employer. This shows that this respondents did not care much about language as long as the immigrant was guaranteed to contribute to the country. Negatively correlated with the above are the immigrants who came to America for jobs but with no Job plan. Immigrants using an interpreter also occupy similar spot. Also, the group of Immigrants escaping Persecution is not really correlated with any other categories.

7 Validation using MCA

Although MCA provides nice data visualization and interpretation, the outputs are difficult to assess. As in, finding the statistical properties of the category factors in MCA can be theoretically hard to find, as the distribution of eigen values is more complex than that of CA, and their sum does not have the same meaning.

In this section, we use *Internal Validation* based on resampling techniques such as bootstrapping to find the confidence ellipsoids of the factors of the variables inn 2-dimension, and see if the sample is a good representative of the population.

Our data is the `wg93` data from **FactomineR** package in R.This dataset is from the International Social Survey Program (ISSP 1993). Each of the 871 respondents were asked the following four questions:

A. We believe too often in science, and not enough in feelings and faith

B. Overall modern science does more harm than good.

C. Any change humans cause in nature-no matter how scientific-is likely to make things worse.

D. Modern Science will solve our environmental problems with little change to our way of life.

Each question has five possible response categories:

1. Agree Strongly
2. Agree
3. Neither agree nor disagree
4. Disagree
5. Disagree strongly

Note that the first three questions pertain to an anti-scientific outlook, whereas the last one is more pro-science.

First we look at the biplot of MCA on this data:

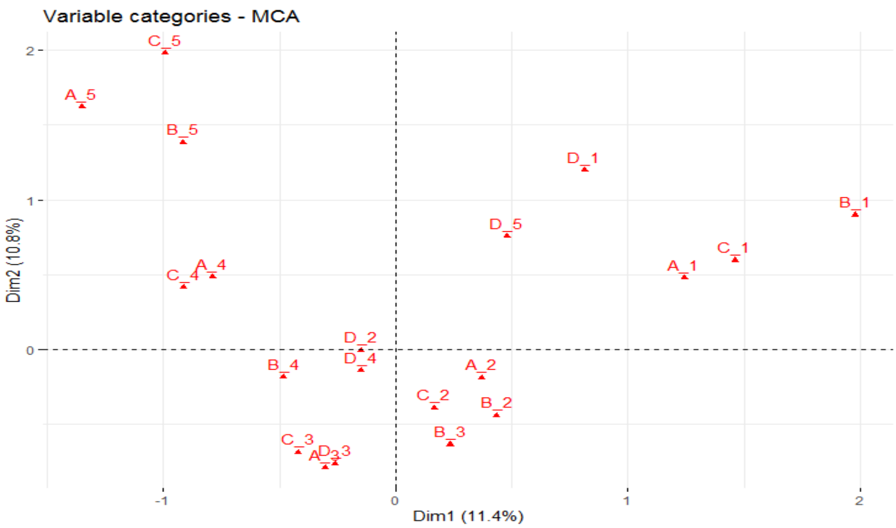


Figure 4: Biplot

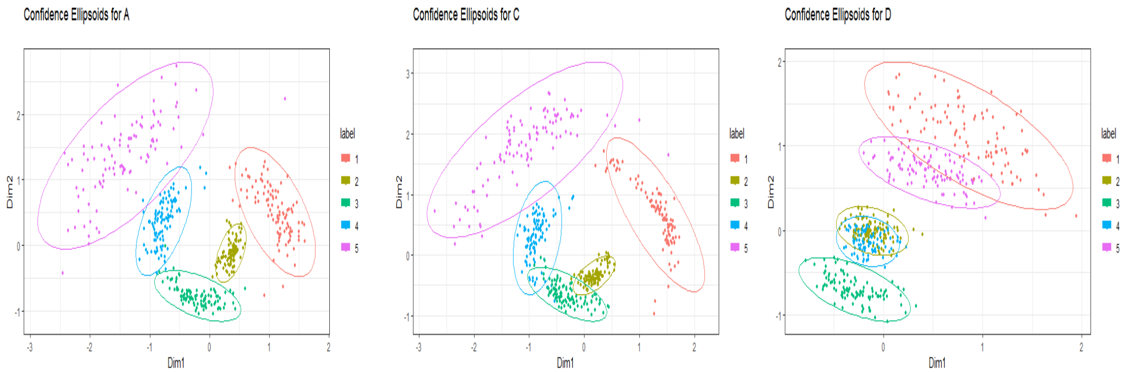
From the biplot, we can see that the first principal axis can be interpreted as one going from "Disagree strongly" to "Agree Strongly". We see that questions A, B and C follow a very similar pattern, with strong disagreements on the left to strong agreements on the right, in a wedge-shaped horseshoe pattern. Question D, however, has a completely dierent trajectory, with the two poles of the scale very close together. Now the rst three questions were all worded negatively towards science whereas question D was worded positively, so we would have expected D5 to lie towards A1, B1 and C1, and D1 on the side of A5, B5 and C5. Such a pattern of response is known as *Guttman Effect* or *Horseshoe Effect*(Van Rijckevorsel 1987). We can do two-step MCA or non-linear PCA to get rid of this. The fact that D1 and D5 lie close together inside the horseshoe means that they are both associated with the extremes of the other three questions — the most likely explanation is that some respondents are misinterpreting the change of direction of the fourth question.

Another possible pattern visible in this biplot is the *Battery Effect*, often observed in survey analysis. When several questions have the same response categories, the respondent often chooses identical answers without sufficiently considering the content of the questions. That could account for the grouping for "Neither Agree nor Disagree" and "Disagree" in the bottom part of the plot.

Now, to get an estimate of how the factor scores fluctuate, and to see if they are stable estimates of the true factor scores (i.e if the relation that we get between different variables via biplot, is stable) we perform *Total Bootstrapping Internal Validation Procedure.*

The total bootstrap is a conservative validation procedure. A specific replication is generated by a drawing with replacement of n individuals from the rows of $\mathbf{X}$. Each replication k leads to a separate MCA. We plot the factor scores of all the MCAs in the same principal plane, corresponding to the principal axes of the original MCA. To remedy the arbitrariness of the signs of the axes in each replication, the orientation of the replicated axes are *a posteriori* changed (where necessary) to maintain a positive correlation with the axes of the original MCA.

We perform $K = 100$ bootstraps, and form the 95% confidence ellipsoids for the levels of A, C and D.



(a) (b) (c)

Figure 5: Bootstrap Confidence Ellipsoids

In the case of A and C (and also B, which is not displayed here), the response categories belonging to a specific question appear to be distinct. Also, the confidence areas of homologous categories in A and C (and B) have similar sizes, with almost all sets of identical categories relating to A,B and C overlap. For example we can see A5 and C5 largely overlap. But for question D, the confidence ellipsoid for "Agree Strongly" and "Disagree Strongly" has a significant intersection, whereas the categories Agree and Disagree almost overlap. This hints at the misinterpretation of this question, and also battery effect for the option "Disagree".

8 Prediction and Classification using MCA

Classifying observations described by Categorical Predictors into one of the k classes can be done using log-linear graphical model, but these model suffer from Curse of Dimensionality and are difficult to apply when the number of categories is large. Saporta(1976) suggested the following for tackling this problem:

1. An MCA is performed on the predictors.
2. A Fisher's Linear Discriminant Analysis is done using factor coordinates as predictors.

This method is kown as *Disqual.*

We use the `Poison` data from the **FactomineR** package for this classification purpose. This data appeared in Box, G. E. P. and D. R. Cox, *An Analysis of Transformations (with discussion),* Journal of the Royal Statistical Society, Series B, Vol. 26, No. 2, pp. 211–254. The data used here refer to a survey carried out on a sample of children of primary school who suffered (and did not suffer) from food poisoning. They were asked about their symptoms and about what they ate. The predictors are:

- Fish.
- Mayo.
- Courgate.
- Cheese.
- Ice-cream.

Each of the predictor have two levels-Yes or No. The response variable is "sick" with categories "Yes" or "No". First we perform an MCA and obtain the individual factor scores. Then we divide the whole dataset into two parts: Training and Test set, and perform a Fisher's LDA on Training set and run it on Test set. We do this bi-partition of data 10 times in a 10-fold cross validation. The confusion matrix is given by:

	Observed Not Sick	Observed Sick
Predicted Not Sick	16.84%	11.58%
Predicted Sick	15.26%	56.32%

Table 1: LDA Confusion Matrix

We see that almost 83% of sick students have been predicted to be sick and almost 52% of not-sick students have been predicted to be not sick. Now, we had in total only 55% data poinnts, and more samples would have resulted in a clearer reflection of thee efficiency of thhe method. Further to improve its efficiency, we can use Logistic Regression/Support Vector Machine. Thus using MCA we can classify the data.

9 Bibliography

1. Jean-Paul Benzécri. *Sur le calcul des taux d'inertie dans l'analyse d'un questionnaire: Addendum et Erratum.*
2. Pierre Bourdieu. *La distinction.*
3. Härdle and Simar. *Applied Multivariate Statistical Analysis.*
4. Greenacre, Michael and Hastie, Trevor. *The Geometric Interpretation of Correspondence Analysis*, Journal of the American Statistical Association, Vol. 82, No. 398 (Jun., 1987), pp.437-447.
5. Greenacre, Michael and Blasius, Jörg (editors) (2006). *Multiple Correspondence Analysis and Related Methods.* London: Chapman & Hall/CRC.
6. Greenacre, Michael (2007). *Correspondence Analysis in Practice*, Second Edition. London: Chapman & Hall/CRC.

$\mathfrak{S}$