SCA vs. TCA: an Expertise

Sergio Camiz and Gastão Coelho Gomes

24/07/2015

1 Introduction

The aim of this work is to compare both theoretically and in practice two exploratory methods whose aim is apparently the same, applied to a two-way contingency table: to represent both rows and column levels on the same graphical (reduced dimensional) space, in order to help interpretability. As interpretability we mean that the relations that exist in the table may be seen graphically in terms of both absolute and relative position of the points-levels. The methods are *Correspondence Analysis* (*SCA*, Benzécri et al., 1973-82; Greenacre, 1983) and *Taxicab Correspondence Analysis* (*TCA*, Choulakian, 2006) with their extensions to multiple tables *Multiple Correspondence Analysis* (*MCA*, Benzécri et al., 1973-82; Greenacre, 1983) and *Taxicab Multiple Correspondence Analysis* (*TMCA*, Choulakian, 2008).

In the following, let $N = (n_{ij})$ an $r \times c$ contingency table, with $n = n_{..}$ its grand total, that is the number of units, $P = (p_{ij}) = (n_{ij}/n)$ the corresponding matrix of relative frequencies, $\boldsymbol{r} = (p_{1.}, ..., p_{r.})'$ the vector of row marginal profile $\boldsymbol{c} = (p_{.1}, ..., p_{.c})'$ the vector of column marginal profile, and $D_r = diag(\boldsymbol{r})$, $D_c = diag(\boldsymbol{c})$ the corresponding diagonal matrices. In the following, we concentrate on matrix P, since n, the number of units, in all formulas is a scale factor and is relevant only in the statistical tests. It is well known that the matrix $\boldsymbol{rc}'$ represents the matrix of independence among the crossing characters, so that we may be only interested to study, and thus to graphically represent, the matrix of deviations from independence $D = P - \boldsymbol{rc}'$.

For this purpose, we must get pairs of unit vectors of coordinates $(\boldsymbol{c}_{\boldsymbol{r}}^{\alpha}, \boldsymbol{c}_{\boldsymbol{c}}^{\alpha})$, for the levels of the characters by row and column, respectively, with $\alpha = 1, \ldots, s = \min(r, c) - 1$, with the requirement of orthogonality. As the graphical representation aims at outlining these deviations, we may wish that these coordinates represent deviations and for that the additive model of data reconstruction is adopted, that is

$$d_{ij} = p_{ij} - p_i.p._j = p_i.p._j \sum_{\alpha=1}^{s} \iota_\alpha {c_r}_i^\alpha {c_c}_j^{\alpha\prime} \tag{1}$$

with the conditions

$$\begin{gathered}\sum_{ij}(p_{ij} - p_{i\cdot}p_{\cdot j}) = 0 \\ \sum_{i} p_{i\cdot}c_{r_i}^{\alpha} = \sum_{j} p_{\cdot j}c_{c_j}^{\alpha} = 0\ \forall\alpha \\ \sum_{ik} p_{i\cdot}p_{k\cdot}c_{r_i}^{\alpha}c_{r_k}^{\alpha} = \sum_{jh} p_{\cdot j}p_{\cdot h}c_{c_j}^{\alpha}c_{c_h}^{\alpha} = \delta_{ij}\ \forall\alpha\end{gathered} \tag{2}$$

The (2) are ordinary identification conditions on the deviations from expectation and on standardized coordinates. Essentially, the rationale of additive models is to decompose the table into independent additive unit-rank components, $P = \boldsymbol{rc}' + \sum_{\alpha} L_{\alpha}$ that here will be named *layers*, each layer

$$L_{\alpha} = \iota_{\alpha}\, p_{i\cdot}\, p_{\cdot j}\, c_{r_i}^{\alpha}\, c_{c_j}^{\alpha\prime}$$

representing an independent component of the deviation from the independence of the original table. Should the coordinates of both rows and columns be correlated with some other character, one may imagine to attribute to its influence the different levels of the characters crossed in the table.

2 The two methods

The two methods under examination adopt two different metrics in their spaces of representation. Consider two points A and B, whose coordinates are $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$, and a vector $\boldsymbol{v}$, whose components are $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$. We define the following metrics:

- L_2 metrics, also known as *Euclidean*, in which the distance between two points A and B is given by $d_2(A, B) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$ and the induced L_2 norm is thus $\parallel \boldsymbol{v} \parallel_2 = \sqrt{\sum_{i=1}^{n}(v_i)^2}$;
- L_1 metrics, also known as *Manhattan, City block*, or *Taxicab*, in which the distance between two points A and B is given by $d_1(A, B) = \sum_{i=1}^{n} |a_i - b_i|$ and the induced norm is thus $\parallel \boldsymbol{v} \parallel_1 = \sum_{i=1}^{n} |v_i|$;
- L_∞ metrics, in which the distance between two points A and B is given by $d_\infty(A, B) = \max_{i\in(1,n)} |a_i - b_i|$ and the induced norm is thus $\parallel \boldsymbol{v} \parallel_\infty = \max_{i\in(1,n)} |v_i|$.

According to the first two metrics, two Correspondence Analyses are defined, in order to study a contingency data table:

1. *Simple Correspondence Analysis* (*SCA*, Benzécri et al., 1973-82; Greenacre, 1983), based on L_2 metrics and the *Generalized Singular Value Decomposition* (*GSVD* Greenacre, 1983; Abdi, 2007);

2. *Taxicab Correspondence Analysis* (*TCA*, Choulakian, 2006), based on L_1 metrics, and the *Taxicab Singular Value Decomposition* (*TSVD*, Choulakian, 2004).

2.1 Singular Value Decompositions

We may ground our further discussion on the well known Singular Value Decomposition (*SVD*, Greenacre, 1983; Abdi, 2007) theorem, that states

Theorem 1 (Singular Value Decomposition) *Any real matrix X may be decomposed as $X = U\Lambda^{1/2}V'$, with Λ the diagonal matrix of the real non-negative eigenvalues of XX', U the orthogonal matrix of the corresponding eigenvectors, and V the matrix of eigenvectors of $X'X$ (with the same eigenvalues), with both constraints $U'U = I$ and $V'V = I$.*

This theorem corresponds to the reconstruction formula of an r-rank matrix

$$x_{ij} = \sum_{\alpha=1}^{r} \sqrt{\lambda_\alpha} \;\; u_{i\alpha} \; v_{j\alpha}$$

on which the Eckart and Young (1936) theorem is based:

Theorem 2 (Eckart and Young) *The s-rank reconstruction of any real matrix X, with $s < r$, the rank of X, once its singular values are sorted in decreasing order,*

$$x_{ij} \approx \sum_{\alpha=1}^{s} \sqrt{\lambda_\alpha} \;\; u_{i\alpha} \; v_{j\alpha} \tag{3}$$

is the best one in the least-squares sense.

? proposes to build the *SVD* solution through a recursive optimization process. Indeed, it consists in finding the first vectors $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ principal component of a matrix X as the solution of the equivalent optimization problems

$$\begin{aligned} \max \parallel X\boldsymbol{u} \parallel_2, \text{ subject to } \parallel \boldsymbol{u} \parallel_2 = 1; \\ \max \parallel X'\boldsymbol{v} \parallel_2, \text{ subject to } \parallel \boldsymbol{v} \parallel_2 = 1. \end{aligned}$$

The solution gives

$$\lambda_1 = \max_{\boldsymbol{u}} \frac{\parallel X\boldsymbol{u} \parallel_2}{\parallel \boldsymbol{u} \parallel_2} = \max_{\boldsymbol{v}} \frac{\parallel X'\boldsymbol{v} \parallel_2}{\parallel \boldsymbol{v} \parallel_2} = \max_{\boldsymbol{u},\boldsymbol{v}} \frac{\boldsymbol{v}'X\boldsymbol{u}}{\parallel \boldsymbol{u} \parallel_2 \parallel \boldsymbol{v} \parallel_2}$$

which is the largest singular value of X. The complete solution results by recursively applying the optimization problem on the residuals. Thus, the reconstruction formula holds:

$$X = \sum_{\alpha=1}^{\min(r,c)} \lambda_\alpha \boldsymbol{v}_\alpha \boldsymbol{u}'_\alpha$$

and it results

$$\sum_{\alpha} \lambda_\alpha^2 = \mathrm{Tr}(X'X).$$

Note that, if we consider the principal coordinates

$$\begin{aligned} \boldsymbol{f}_\alpha = X\boldsymbol{u}_\alpha, &\text{ with } \boldsymbol{v}'_\alpha \boldsymbol{f}_\alpha = \| \boldsymbol{f}_\alpha \|_2 = \lambda_\alpha \\ \boldsymbol{g}_\alpha = X'\boldsymbol{v}_\alpha, &\text{ with } \boldsymbol{u}'_\alpha \boldsymbol{g}_\alpha = \| \boldsymbol{g}_\alpha \|_2 = \lambda_\alpha \end{aligned}$$

the reconstruction formula becomes

$$X = \sum_{\alpha=1}^{\min(r,c)} \frac{1}{\lambda_\alpha} \boldsymbol{f}_\alpha \boldsymbol{g}'_\alpha$$

Correspondence analysis requires a special metrics, thus we shall refer to Generalized Singular Value Decomposition (*GSVD*, Greenacre, 1983; Abdi, 2007). For a given matrix X, this involves using two positive definite square matrices expressing constraints imposed on both rows and columns of X respectively. If M_r and M_c are such matrices, the $GSVD$ aims at decomposing X as $X = U\Lambda^{1/2}V'$, under the orthogonality constraints $U'M_rU = I$ and $V'M_cV = I$. We shall express these conditions by saying that U and V are required to be M_r- and M_c-orthogonal, respectively.

Theorem 3 (Generalized Singular Value Decomposition) *Given two real positive definite matrices M_r and M_c, any real matrix X may be decomposed as $X = F\Lambda^{1/2}G'$, under constraints $F'MF = I$ and $G'NG = I$.*

The solution is given by the SVD of the matrix $\widetilde{X} = M_r^{1/2} X M_c^{1/2} = U\Lambda^{1/2}V'$, with $U'U = I$, $V'V = I$, $F = M_r^{-1/2}U$, and $G = M_c^{-1/2}V$. It results that $FF' = M_r^{-1}$ and $GG' = M_c^{-1}$ respectively, that is $F'M_rF = I$ and $G'M_cG = I$: thus, we say that F and G are M_r- and M_c-orthogonal, respectively.

Taxicab Singular Value Decomposition In analogy with what proposed for SVD, Choulakian (2004) proposes a recursive method in the Taxicab metrics too. The first vectors are the solution of the equivalent optimization problems

$$\begin{aligned} &\max \| X\boldsymbol{u} \|_1, \text{ subject to } \| \boldsymbol{u} \|_\infty = 1; \\ &\max \| X'\boldsymbol{v} \|_1, \text{ subject to } \| \boldsymbol{v} \|_\infty = 1. \end{aligned}$$

The solution

$$\lambda_1 = \max_{\boldsymbol{u}} \frac{\| X\boldsymbol{u} \|_1}{\| \boldsymbol{u} \|_\infty} = \max_{\boldsymbol{v}} \frac{\| X'\boldsymbol{v} \|_1}{\| \boldsymbol{v} \|_\infty} = \max_{\boldsymbol{u},\boldsymbol{v}} \frac{\boldsymbol{v}'X\boldsymbol{u}}{\| \boldsymbol{u} \|_\infty \| \boldsymbol{v} \|_\infty}$$

is a combinatorial problem described by **?**. The complete solution results by recursively applying the optimization problem on the residuals, but it may be seen as a *TSVD*, *Taxicab Singular Value Decomposition.* The corresponding principal coordinates are

$$\begin{aligned} \boldsymbol{f}_\alpha = X\boldsymbol{u}_\alpha, &\text{ with } \boldsymbol{v}'_\alpha \boldsymbol{f}_\alpha = \| \boldsymbol{f}_\alpha \|_1 = \lambda_\alpha \\ \boldsymbol{g}_\alpha = X'\boldsymbol{v}_\alpha, &\text{ with } \boldsymbol{u}'_\alpha \boldsymbol{g}_\alpha = \| \boldsymbol{g}_\alpha \|_1 = \lambda_\alpha \end{aligned}$$

In this case, since both $\boldsymbol{u}_\alpha$ and $\boldsymbol{v}_\alpha$ are essentially vectors of signs ($\boldsymbol{u}_\alpha = \operatorname{sgn}(\boldsymbol{g}_\alpha)$ and $\boldsymbol{v}_\alpha = \operatorname{sgn}(\boldsymbol{f}_\alpha)$), the reconstruction formula becomes:

$$X = \sum_{\alpha=1}^{\min(r,c)} \frac{1}{\lambda_\alpha} \boldsymbol{f}_\alpha \boldsymbol{g}'_\alpha$$

Note that in L_1 metrics, the total inertia should be the sum of each layer's ones.

2.2 Simple Correspondence Analysis

Correspondence Analysis may be formulated according to different points of view. We try to ground it on *SVD*. We know that the relations between rows and columns of N are summarized by the χ^2 statistics, that measures the departure from the independence between rows and columns. Since the independence is estimated by $N_0 = nP_0 = n\boldsymbol{r}\boldsymbol{c}'$, the departure from independence is estimated by

$$\chi^2 = n\ \phi^2 = n\ \sum_i \sum_j \frac{(p_{ij} - p_i.p._j)^2}{p_i.p._j} \tag{4}$$

with $(r-1) \times (c-1)$ degrees of freedom. Note that N and its grand total n are interesting only to evaluate the chi-square significance, so that interest may be concentrated most on the matrix P. Note that, by simplifying (4), ϕ^2 may be computed directly as

$$\phi^2 = \sum \frac{p_{ij}^2}{p_i.p._j} - 1. \tag{5}$$

We may compute both in an alternative way: (5) may be written as

$$\phi^2 = \operatorname{trace}(S'S) - 1 \text{ with } S = \frac{p_{ij}}{\sqrt{p_i.p._j}}$$

and (4), may be written as

$$n\ \operatorname{trace}(\dot{S}'\dot{S}) = n\ \operatorname{trace}\left(\left(\frac{p_{ij} - p_i.p._j}{\sqrt{p_i.p._j}}\right)' \left(\frac{p_{ij} - p_i.p._j}{\sqrt{p_i.p._j}}\right)\right)$$

that is, in matrix form

$$\phi^2 = \operatorname{trace}\left((P - \boldsymbol{r}\boldsymbol{c}')' D_r^{-1} (P - \boldsymbol{r}\boldsymbol{c}') D_c^{-1}\right) \tag{6}$$

We refer here to the possibility to partition the chi-square into components. Indeed, if we succeed in writing N as sum of independent tables, we may partition the chi-square accordingly and check for significance of each component independently. Our problem is to reduce the rank of P (and consequently of N) without losing relevant information. Indeed, we may formalize the problem, considering a suitable reduced rank matrix $\hat{P}$ that best approximates P in the sense of the weighed least squares, that is minimizing the residuals:

$$R = n \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(p_{ij} - \hat{p}_{ij})^2}{p_{i\cdot} p_{\cdot j}} = n \text{ trace} \left((P - \hat{P})' D_r^{-1} (P - \hat{P}) D_c^{-1} \right) \tag{7}$$

where the weights are the inverse of the expected frequencies. Note that this formulation allows to check for significance of the residuals, since R may be tested as a chi-square with ??? degrees of freedom.

For this purpose, we may apply the SVD to $\widetilde{P} = D_r^{-1/2} P D_c^{-1/2} = U \Lambda^{1/2} V'$, with $U'U = I$, $V'V = I$. This corresponds to apply $GSVD$ to the table P with the constraints given by the diagonal matrices D_r^{-1} and D_c^{-1}, that is, by decomposing $P- = D_r^{1/2} U \Lambda^{1/2} V' D_c^{1/2} = F \Lambda^{1/2} G'$, with $F = D_r^{1/2} U$, and $G = D_c^{1/2} V$, with $FF' = D_r$ and $GG' = D_c$,, that is F and G D_r- and D_c-orthogonal. Thus, the reconstruction formula may be well synthesized as

$$N = nP = nD_r U \Lambda^{1/2} V' D_c = nF \Lambda^{1/2} G'. \tag{8}$$

with the best reduced rank approximations based on the Eckart-Young theorem: for any $q \leq rank(P) \leq \min(r, c)$, the partial q-rank reconstruction formula (3) becomes:

$$n_{ij} \approx \hat{n}_{ij,q} = n \ \hat{p}_{ij,q} = n \ p_{i\cdot} p_{\cdot j} \left(\sum_{\alpha=1}^{q} \sqrt{\lambda_\alpha} \ u_{i\alpha} \ v_{j\alpha} \right) = n \left(\sum_{\alpha=1}^{q} \sqrt{\lambda_\alpha} \ f_{i\alpha} \ g_{j\alpha} \right).$$

where the equality holds for $q = rank(P)$.

Thus, F and G provide factors D_r- and D_c-orthogonal respectively, whereas we are interested in getting coordinates whose weighed inertia sums to the corresponding eigenvalue. To get this, we define $\Phi = D_r^{-1/2} U \Lambda^{1/2}$ and $\Psi = D_c^{-1/2} V \Lambda^{1/2}$, so that

$$\Phi' D_r \Phi = \Lambda = \Psi' D_c \Psi. \tag{9}$$

As $F = D_r \Phi \Lambda^{-1/2}$ and $G = D_c \Psi \Lambda^{-1/2}$, if we introduce these transformations into (8) we get:

$$N = n \ P = n \ D_r \Phi \Lambda^{-1/2} \Lambda^{1/2} \Lambda^{-1/2} \Psi D_c = n \ D_r \Phi \Lambda^{-1/2} \Psi D_c. \tag{10}$$

Consequently, the partial q-rank reconstruction formula becomes:

$$n_{ij} \approx \hat{n}_{ij,q} = n \ \hat{p}_{ij,q} = n \ r_i c_j \left(\sum_{\alpha=1}^{q} \frac{1}{\sqrt{\lambda_\alpha}} \ \phi_{i\alpha} \ \psi_{j\alpha} \right).$$

It is well known that the first eigenvalue equals 1 and the corresponding eigenvectors are the marginals. Thus, one may write

$$n_{ij} \approx \hat{n}_{ij,q} = n \ \hat{p}_{ij,q} = n \ r_i c_j \left(1 + \sum_{\alpha=2}^{q} \frac{1}{\sqrt{\lambda_\alpha}} \ \phi_{i\alpha} \ \psi_{j\alpha} \right).$$

An alternative is to consider directly the deviation from the expectation under independence $D = P - \boldsymbol{r}\boldsymbol{c}'$. This leads to the same reconstruction, that is

$$n_{ij} \approx \hat{n}_{ij,q} = n\ \hat{p}_{ij,q} = n\ r_i c_j \left(1 + \sum_{\alpha=1}^{q-1} \frac{1}{\sqrt{\lambda_\alpha}}\ \phi_{i\alpha}\ \psi_{j\alpha}\right). \tag{11}$$

It is customary to consider the layers' *inertia* ι_α^2 as a measure of their importance, so that they are usually sorted in decreasing inertia order. Indeed, on the statistical point of view, since inertias along each dimension α equal $\phi_\alpha^2 = \iota_\alpha^2$, a partial chi-square may be associated ${\chi_\alpha}^2 = n\phi_\alpha^2 = n\iota_\alpha^2$ that may be tested against independence with $df = (r + c - 2\alpha - 1)$ (Kendall and Stuart, 1961; Orlóci, 1978). Note in addition that, as inertias sum up to the table ϕ^2, the total chi-square results

$$\chi^2 = n\phi^2 = n \sum_\alpha \phi_\alpha^2 = n \sum_\alpha \iota_\alpha^2.$$

It is possible to select layers that contain a significant deviation from the independence. Indeed, given a partial reconstruction of the original table limited to the first $r < s$ layers, the classical test for goodness of fit (Kendall and Stuart, 1961) may be applied, or more easily the Malinvaud (1987) test. The test may be applied, as, for each α-dimensional partial reconstruction, the residuals correspond to

$$Q_\alpha = \sum_{ij} \frac{\left(n_{ij} - \widetilde{n}_{\alpha ij}\right)^2}{\widetilde{n}_{\alpha ij}},$$

asymptotically chi-square-distributed with $(r - \alpha - 1) \times (c - \alpha - 1)$ degrees of freedom. In the formula, $\widetilde{n}_{\alpha ij}$ is the cell value estimated by the α-dimensional solution, and the table chi-square test results when $\alpha = 0$ and $\widetilde{n}_{0ij} = \frac{n_{i.}\ n_{.j}}{n_{..}}$ is the expected value under independence. Now, Malinvaud (1987) showed that, by substituting the estimated cell values with the expected ones under independence hypothesis, the formula may be approximated by

$$\widetilde{Q}_\alpha = \sum_{ij} \frac{\left(n_{ij}\ - \widetilde{n}_{\alpha ij}\right)^2}{n r_i c_j}\ = \chi^2 - \sum_{\beta=1}^{\alpha} \chi_\beta^2 = n \sum_{\gamma=\alpha+1}^{s} \iota_\gamma^2,$$

that may be more easily used to check for nullity of the residuals. Opposite to the individual layer's test above, Malinvaud's is an overall one, that may be used to reject the hypothesis of the residuals randomness.

2.3 Taxicab Correspondence Analysis

Taxicab Correspondence Analysis is defined as the Taxicab Singular Value Decomposition of the data table $D = P - \boldsymbol{r}\boldsymbol{c}'$, taking into account the table's *profiles*, respectively $R = D_r^{-1} D$ for the rows and $C = D_c^{-1} D$ for the columns.

Unlike *SCA*, the solution is recursive, considering at each step the residuals from the previous factors. This leads to the reconstruction formula

$$P = \boldsymbol{p}_r \boldsymbol{p}_c' + \sum_{\alpha=2}^{\min(r,c)} \frac{1}{\lambda_\alpha} \boldsymbol{f}_\alpha \, \boldsymbol{g}_\alpha'.$$

since the first factor is shown to correspond to the independence, with λ_α the L_1-measure of dispersion along the α-th factor (note that $\lambda_1 = 1$). Expressed elementwise the formula becomes:

$$p_{ij} = p_{i.} p_{.j} + \sum_{\alpha=2}^{\min(r,c)} \frac{1}{\lambda_\alpha} f_{i\alpha} \, g_{j\alpha}.$$

Now, if we transform the coordinates $F_{i\alpha} = \frac{f_{i\alpha}}{p_{i.}}$ and $G_{j\alpha} = \frac{f_{i\alpha}}{p_{.j}}$ we get

$$n_{ij} = n \; r_i c_j \left(1 + \sum_{\alpha=2}^{\min(r,c)} \frac{1}{\lambda_\alpha} F_{i\alpha} \, G_{j\alpha} \right). \tag{12}$$

just as for *SCA*.

2.4 Multiple Correspondence Analysis and Taxicab

It is well known that *MCA* is defined as the *SCA* of an indicator matrix Z, describing the levels of several nominal characters. Indeed, it may also be done by applying *SCA* to the Burt's matrix $Z'Z$, a super-table that crosses all characters producing the corresponding contingency tables. Unlike the L_2 analysis, *Taxicab Multiple Correspondence Analysis* (*TMCA*) produces different results depending on which table the analysis is run.

3 An example: the Snee data

As an example, we take the Snee (1978) data table that crosses 592 students of the University of Delaware according to the color of the eyes and of the hair, both with 4 levels. The table N is thus:

	Hair				
Eyes	Black	Brown	Red	Blond	Total
Dark Brown	68	119	26	7	220
Light Brown	15	54	14	10	93
Green	5	29	14	16	64
Blue	20	84	17	94	215
Total	108	286	71	127	592

We know that the table under the hypothesis of independence is given by the product of the marginals $\boldsymbol{r}$ and $\boldsymbol{c}$, that is:

```
> r=apply(snee,1,sum); r
 Dark Brown Light Brown       Green        Blue
        220          93          64         215
> c=apply(snee,2,sum); c
Black Brown   Red Blond
  108   286    71   127

> r\%*\%t(c)/sum(snee)
                   Hair
   Eyes            Black     Brown       Red     Blond Total
   Dark Brown  40.13514 106.28378 26.385135  47.19595   220
   Light Brown 16.96622  44.92905 11.153716  19.95101    93
   Green       11.67568  30.91892  7.675676  13.72973    64
   Blue        39.22297 103.86824 25.785473  46.12331   215
   Total      108        286        71        127         592
```

We may apply *CA* from the *R* package *FactoMineR*.

```
library(FactoMineR)
cs<-CA(snee); summary(cs)
Call:
CA(snee)

Eigenvalues   Dim.1   Dim.2   Dim.3
Variance      0.209   0.022   0.003
\% of var.   89.373   9.515   1.112
Cumul. \%    89.373  98.888 100.000


Rows            Dim.1     ctr    cos2     Dim.2     ctr    cos2     Dim.3     ctr    cos2
Dark Brown  |  -0.492  43.116   0.967 |  -0.088  13.042   0.031 |   0.022   6.680   0.002 |
Light Brown |  -0.213   3.401   0.542 |   0.167  19.804   0.336 |  -0.101  61.086   0.121 |
Green       |   0.162   1.355   0.176 |   0.339  55.910   0.773 |   0.088  31.925   0.052 |
Blue        |   0.547  52.128   0.977 |  -0.083  11.244   0.022 |  -0.005   0.310   0.000 |

Columns         Dim.1     ctr    cos2     Dim.2     ctr    cos2     Dim.3     ctr    cos2
Black       |  -0.505  22.246   0.838 |  -0.215  37.877   0.152 |   0.056  21.633   0.010 |
Brown       |  -0.148   5.086   0.864 |   0.033   2.319   0.042 |  -0.049  44.284   0.094 |
Red         |  -0.130   0.964   0.133 |   0.320  55.131   0.812 |   0.083  31.913   0.055 |
Blond       |   0.835  71.704   0.993 |  -0.070   4.673   0.007 |   0.016   2.171   0.000 |
```

In the following are reported the statistics concerning the significance of the table and of the eigenvectors:

```
[1] "Partition of data table chi-square."
[1] "Number of rows     = " "4"
[1] "Number of columns  = " "4"
[1] "Grand total        = " "592"
[1] "Maximum dimension  = " "3"
[1] "Degrees of freedom = " "9"
[1] "Trace (sum of eig.)= " "0.233597705449338"
[1] "Eigenvalues > 0.0  = " "3"
[1] "Total chi-square   = " "138.289841626008"
[1] "Probability        = " "0"
[1] "Test value         = " "Inf"
     N         eig          %      Cum%       CorC        Chi df       p-val    v-test          Res df       p-val    v-test
[1,] 1 0.208772652 0.89372732 0.8937273 0.45691646 123.593410  5 0.000000000       Inf 1.469643e+01  4 0.005374079 2.5507818
[2,] 2 0.022226615 0.09514911 1.7874546 0.14908593  13.158156  3 0.004306773 2.6270232 1.538276e+00  1 0.214874599 0.7896209
[3,] 3 0.002598439 0.01112356 1.8826038 0.05097489   1.538276  1 0.214874599 0.7896209 4.884981e-15  0 0.000000000       Inf
```

It results that the table is significant and so are the first two eigenvectors. Note also that the first factor canonical correlation is .45, a medium value.

Here, the coordinates are such that their weighed average $\sum_i p_i.c_{\alpha i} = \sum_j p_{\cdot j} c_{\alpha j}$ is zero and the sum of squares equals the corresponding eigenvalue: $\sum_i p_i.c^2_{\alpha i} = \sum_j p_{\cdot j} c^2_{\alpha j} = \lambda_\alpha$. Indeed:

```
> sum(cs$row$coord[,1]*r) [1]  1.00614e-16   > sum(cs$row$coord[,1]^2*r) [1] 0.2087727
> sum(cs$row$coord[,2]*r) [1]  2.428613e-17  > sum(cs$row$coord[,2]^2*r) [1] 0.02222661
> sum(cs$row$coord[,3]*r) [1] -1.431147e-17  > sum(cs$row$coord[,3]^2*r) [1] 0.002598439
> sum(cs$col$coord[,1]*c) [1] -1.734723e-17  > sum(cs$col$coord[,1]^2*c) [1] 0.2087727
> sum(cs$col$coord[,2]*c) [1] -6.418477e-17  > sum(cs$col$coord[,2]^2*c) [1] 0.02222661
> sum(cs$col$coord[,3]*c) [1]  3.339343e-17  > sum(cs$col$coord[,3]^2*c) [1] 0.002598439
```

Now, applying the reconstruction formula (11) we get the independence table and the three following layers that sum up to the table:

```
L0 = n*r%*%t(c)
rownames(L0) = rownames(st); L0

L0              Black     Brown       Red    Blond
Dark Brown   40.13514 106.28378 26.385135 47.19595
Light Brown 16.96622  44.92905 11.153716 19.95101
Green       11.67568  30.91892  7.675676 13.72973
Blue        39.22297 103.86824 25.785473 46.12331

L1 = L0 * (cs$row$coord[,1] %*% t(cs$col$coord[,1])) / sqrt(cs$eig[1,1]); L1

L1                Black      Brown       Red      Blond
Dark Brown   21.812583  16.972159  3.681074 -42.465815
Light Brown   3.983090   3.099203  0.672183  -7.754476
Green        -2.085516  -1.622720 -0.351950   4.060186
Blue        -23.710157 -18.448642 -4.001307  46.160106

L2 = L0 * (cs$row$coord[,2] %*% t(cs$col$coord[,2])) / sqrt(cs$eig[2,1]); L2

L2               Black     Brown       Red     Blond
Dark Brown   5.107757 -2.056825 -4.996358  1.945426
Light Brown -4.092195  1.647872  4.002945 -1.558622
Green       -5.703893  2.296881  5.579493 -2.172481
Blue         4.688331 -1.887928 -4.586080  1.785677

L3 = L0 * (cs$row$coord[,3] %*% t(cs$col$coord[,3])) / sqrt(cs$eig[3,1]); L3

L3                Black      Brown        Red       Blond
Dark Brown   0.9445252 -2.1991170  0.9301488  0.32444307
Light Brown -1.8571111  4.3238705 -1.8288444 -0.63791503
Green        1.1137334 -2.5930807  1.0967815  0.38256584
Blue        -0.2011475  0.4683273 -0.1980859 -0.06909388

> L0+L1
              Black     Brown       Red    Blond
Dark Brown  61.94772 123.25594 30.066209  4.73013
Light Brown 20.94931  48.02826 11.825899 12.19654
Green        9.59016  29.29620  7.323726 17.78992
Blue        15.51282  85.41960 21.784166 92.28342

> L0+L1+L2
               Black     Brown      Red     Blond
Dark Brown  67.055475 121.19912 25.06985  6.675557
Light Brown 16.857111  49.67613 15.82884 10.637915
Green        3.886267  31.59308 12.90322 15.617434
Blue        20.201148  83.53167 17.19809 94.069094

> L0+L1+L2+L3
            Black Brown Red Blond
Dark Brown     68   119  26     7
Light Brown    15    54  14    10
Green           5    29  14    16
Blue           20    84  17    94
```

We may apply the TCA through the R package TCA to the same table and we obtain:

```
library(TCA)
Ts=TCA(snee,Naxes=3,Graph=TRUE)
Ts

$VectMax
     Axe_1      Axe_2      Axe_3
0.33883081 0.08519358 0.03510355

$A
                  Axe_1       Axe_2        Axe_3
Dark Brown  -0.135797115 -0.02400496  0.008775888
Light Brown -0.033618289  0.02400496 -0.008775888
Green        0.007669832  0.01859184  0.008775888
Blue         0.161745572 -0.01859184 -0.008775888

$B
             Axe_1         Axe_2         Axe_3
Black -0.087495435 -4.259679e-02 -6.938894e-18
Brown -0.073605278  1.288852e-02 -1.755178e-02
Red   -0.008314691  2.970827e-02  1.755178e-02
Blond  0.169415404 -1.821460e-17  1.416520e-17

$F
                 Axe_1       Axe_2       Axe_3
Dark Brown  -0.36541769 -0.06459515  0.02361512
Light Brown -0.21400029  0.15280574 -0.05586372
Green        0.07094595  0.17197448  0.08117697
Blue         0.44536455 -0.05119240 -0.02416431

$G
           Axe_1         Axe_2         Axe_3
Black -0.47960460 -2.334935e-01 -3.803542e-17
Brown -0.15235778  2.667834e-02 -3.633095e-02
Red   -0.06932813  2.477084e-01  1.463472e-01
Blond  0.78971590 -8.490584e-17  6.602989e-17
```

The matrices F and G contain the coordinates that are centered and whose L_1-norm equals the one in $VectMax$:

```
> sum(r%*%sr[,1])  [1] -4.263256e-14    > sum(r%*%abs(sr[,1]))/sum(r)  [1] 0.3388308
> sum(r%*%sr[,2])  [1] -8.881784e-15    > sum(r%*%abs(sr[,2]))/sum(r)  [1] 0.08519358
> sum(r%*%sr[,3])  [1] 8.881784e-15     > sum(r%*%abs(sr[,3]))/sum(r)  [1] 0.03510355
> sum(c%*%sc[,1])  [1] -1.421085e-14    > sum(c%*%abs(sc[,1]))/sum(c)  [1] 0.3388308
> sum(c%*%sc[,2])  [1] -2.49939e-14     > sum(c%*%abs(sc[,2]))/sum(c)  [1] 0.08519358
> sum(c%*%sc[,3])  [1] 1.016215e-14     > sum(c%*%abs(sc[,3]))/sum(c)  [1] 0.03510355
```

As well, we apply here the reconstruction formula **??** and we obtain the three following layers that sum up to the table:

```
ss = Ts$VectMax; ss   # L1 inertias
sr = Ts$F; sr         # row coordinates
sc = Ts$G; sc         # col coordinates

LT1=L0*((sr[,1] %*% t(sc[,1])) / (ss[1])); LT1
L0+LT1
snee-(L0+LT1)
LT2=L0*((sr[,2] %*% t(sc[,2])) / (ss[2])); LT2
L0+LT1+LT2
snee-(L0+LT1+LT2)
LT3=L0*((sr[,3] %*% t(sc[,3])) / (ss[3])); LT3
L0+LT1+LT2+LT3
snee-(L0+LT1+LT2+LT3)

> ss = Ts$VectMax; ss   # singular values or eigenvalues?
     Axe_1      Axe_2      Axe_3
```

```
0.33883081 0.08519358 0.03510355
>
> sr = Ts$F; sr         # row coordinates
                  Axe_1       Axe_2       Axe_3
Dark Brown  -0.36541769 -0.06459515  0.02361512
Light Brown -0.21400029  0.15280574 -0.05586372
Green        0.07094595  0.17197448  0.08117697
Blue         0.44536455 -0.05119240 -0.02416431
>
> sc = Ts$G; sc         # col coordinates
            Axe_1         Axe_2         Axe_3
Black -0.47960460 -2.334935e-01 -3.803542e-17
Brown -0.15235778  2.667834e-02 -3.633095e-02
Red   -0.06932813  2.477084e-01  1.463472e-01
Blond  0.78971590 -8.490584e-17  6.602989e-17

> LT1=L0*((sr[,1] %*% t(sc[,1])) / (ss[1])); LT1
                 Black       Brown       Red      Blond
Dark Brown   20.759398  17.4637825  1.972766 -40.195946
Light Brown   5.139251   4.3233797  0.488383  -9.951014
Green        -1.172492  -0.9863558 -0.111422   2.270270
Blue        -24.726156 -20.8008063 -2.349727  47.876689
>
> L0+LT1
               Black     Brown       Red Blond
Dark Brown  60.89453 123.74757 28.357901     7
Light Brown 22.10547  49.25243 11.642099    10
Green       10.50318  29.93256  7.564254    16
Blue        14.49682  83.06744 23.435746    94
>
> snee-(L0+LT1)
                Black      Brown       Red         Blond
Dark Brown   7.105467 -4.7475663 -2.357901  1.421085e-14
Light Brown -7.105467  4.7475663  2.357901 -1.776357e-15
Green       -5.503183 -0.9325631  6.435746  0.000000e+00
Blue         5.503183  0.9325631 -6.435746  0.000000e+00
>
> LT2=L0*((sr[,2] %*% t(sc[,2])) / (ss[2])); LT2
                Black     Brown       Red         Blond
Dark Brown   7.105467 -2.149903 -4.955564  3.038332e-15
Light Brown -7.105467  2.149903  4.955564 -3.038332e-15
Green       -5.503183  1.665100  3.838083 -2.353188e-15
Blue         5.503183 -1.665100 -3.838083  2.353188e-15
>
> L0+LT1+LT2
            Black     Brown      Red Blond
Dark Brown     68 121.59766 23.40234     7
Light Brown    15  51.40234 16.59766    10
Green           5  31.59766 11.40234    16
Blue           20  81.40234 19.59766    94
>
> snee-(L0+LT1+LT2)
                    Black     Brown       Red        Blond
Dark Brown  -2.842171e-14 -2.597663  2.597663 1.154632e-14
Light Brown -3.552714e-15  2.597663 -2.597663 1.776357e-15
Green       -3.552714e-15 -2.597663  2.597663 1.776357e-15
Blue        -7.105427e-15  2.597663 -2.597663 0.000000e+00
>
> LT3=L0*((sr[,3] %*% t(sc[,3])) / (ss[3])); LT3
                    Black     Brown       Red         Blond
Dark Brown  -1.026956e-15 -2.597663  2.597663  2.096449e-15
Light Brown  1.026956e-15  2.597663 -2.597663 -2.096449e-15
Green       -1.026956e-15 -2.597663  2.597663  2.096449e-15
Blue         1.026956e-15  2.597663 -2.597663 -2.096449e-15
>
> L0+LT1+LT2+LT3
            Black Brown Red Blond
Dark Brown     68   119  26     7
```

```
Light Brown    15    54  14    10
Green           5    29  14    16
Blue           20    84  17    94
>
> snee-(L0+LT1+LT2+LT3)
                    Black Brown           Red        Blond
Dark Brown  -2.842171e-14     0 -1.065814e-14 9.769963e-15
Light Brown -5.329071e-15     0 -3.552714e-15 3.552714e-15
Green       -2.664535e-15     0  0.000000e+00 0.000000e+00
Blue        -7.105427e-15     0  0.000000e+00 0.000000e+00
```

In Figure 1 are shown the scatter plots of both characters labels on the planes spanned by the first two factors of *SCA* (Figure 1 *left* and *right*, respectively).

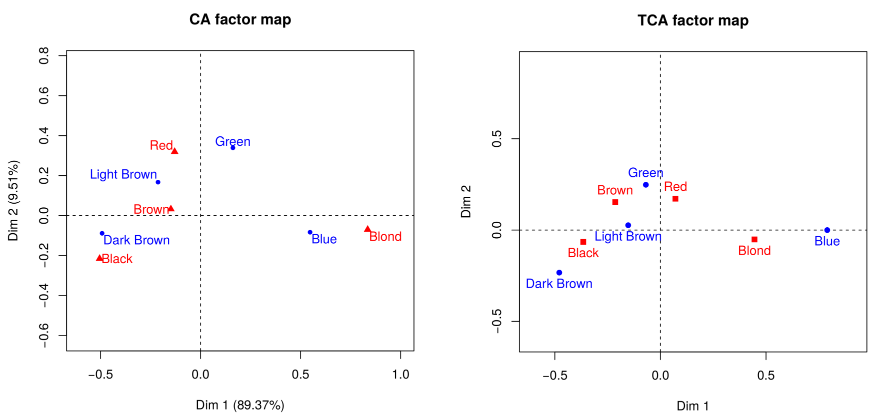

Figure 1: *The scatter plot of both hair and eye colours, according to Snee (1978), on the first factor plane issued by FactoMineR's CA correspondence analysis method (left) and that issued by TCA taxicab method (right).*

4 Another example: the "Palavras" data

We consider a three-way data table taken from Nardy (2007) and used already the authors to study *MCA* and its improvements (Camiz and Gomes, 2013). Nardy (2007) study concerns the architecture of grammar proposed in the Distributed Morphology framework (Halle and Marantz, 1993, 1994) to analyze the internal structure of words in Brazilian publications. In particular, that work studies the writers' control over the degree of complexity of their wording, according to the type of texts they are producing. Four types of texts were distinguished: 1) books for children (in the following labeled *T Child*); 2) gossip, fashion, local news (review, *T Revi*); 3) editorials, articles about science for laymen (Divulgation, *T Divu*); and 4) abstracts of academic articles (Summary, *T Summ*). 2000 word-tokens were extracted (500 from each text type), avoiding repetitions; as well, conjugation or declination were not taken into account, because their cause of variation does not affect the word's meaning. The tokens were analyzed according to their grammatical kind, say kind of words (*W Verb, W Noun, W Adj*, the latter for adjective), and the number of internal layers (Two-, *2-Syl*, Three-, *3-Syl*, four and more layers, *4-Syl*), as a measure of the word's complexity. The syntactic criteria for the word decomposition to count the internal layers are discussed by Nardy (2007): in practice, starting from the root, a full word is obtained by adding some endings to the root, that allow to categorize it as a noun, a verb, or an adjective. In this way, the full word is understood as such. Since from a noun a verb may derive, and a noun or an adjective from a verb, etc. several endings may be added, thus raising the number of layers that sum up to a word. As an example, consider three Portuguese words: the first, rosa is a noun (rose), with only two layers: the root ros and the noun ending a; the second, furar is a verb (to make a hole) composed by the root fur, the noun categorizer a, and the verb categorizer r; the third, salinização is a noun (salinization) composed by five layers.

Table 1: *The contingency three-way data table of "palavras" taken from Nardy (2007), referring to 2000 words characterized by type of text, type of word, and number of levels.*

Type of Text	*Type of Words N. of Levels*	*Names*	*Verbs*	*Adjectives*
	2 Sylabes	203	167	63
Childish	*3 Sylabes*	26	6	32
	4 Sylabes	0	1	2
	2 Sylabes	218	126	41
Review	*3 Sylabes*	51	4	27
	4 Sylabes	15	3	11
	2 Sylabes	207	118	74
Divulgation	*3 Sylabes*	51	6	29
	4 Sylabes	15	1	5
	2 Sylabes	160	72	63
Summary	*3 Sylabes*	75	7	61
	4 Sylabes	32	4	74

The run of *MCA* through the *R* package *FactoMineR* gave the following summary results:

```
                      Dim 1   Dim 2   Dim 3   Dim 4   Dim 5   Dim 6   Dim 7
Variance              0.490   0.364   0.343   0.330   0.308   0.273   0.225
% of var.            20.982  15.599  14.718  14.142  13.216  11.692   9.651
Cumulative % of var. 20.982  36.581  51.298  65.440  78.656  90.349 100.000

$coord
```

```
          Dim 1       Dim 2       Dim 3       Dim 4       Dim 5       Dim 6         Dim 7
2 Syl   -0.43305968 -0.03500517  0.02636409  0.07305786 -0.05149101 -0.04231701 -0.3514408994
3 Syl    1.24048516  0.57792833 -0.81985485 -0.77794032  0.21565077 -0.25507680  1.0351841330
4 Syl    1.67791415 -1.44951598  2.36799168  1.60410742 -0.02667821  1.41271789  1.2671202660
W Adj    1.01578986  0.95780556  0.16006550  0.12807384  0.86127594  0.56831548 -0.7573681475
W Noun   0.07644207 -0.58334686 -0.37067078 -0.17214659 -0.60408843  0.14575379  0.0001114993
W Verb  -1.00837810  0.38930532  0.62362725  0.24454848  0.51268721 -0.77473985  0.6350788951
T Child -0.69339914  0.96932279  0.34421092 -0.43930458 -0.69046958  0.84512588  0.2777684178
T Divu  -0.11108040 -0.03983846 -1.05991655  1.30785136  0.27384193  0.05255584  0.1642967135
T Revi  -0.20863375 -1.01679529  0.12744457 -0.92294327  1.03020161  0.14425407 -0.0683108809
T Summ   1.01684456  0.07997456  0.60441729  0.03144450 -0.61106270 -1.04559481 -0.3777834318
```

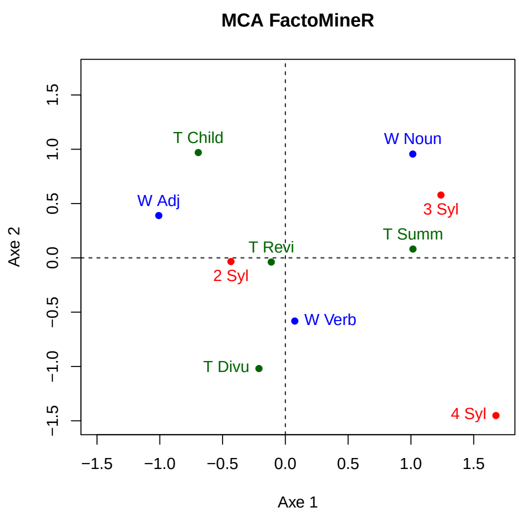

Figure 2: *The scatter plot of levels of Nardy (2007) palavras data, on the first factor plane issued by running Multiple Correspondence Analysis.*

In Figure 2 the pattern of the levels of the three characters is shown on the first factor plane. Note that only the first factor is significant, according to the Ben Ammou and Saporta (1998, 2003) test.

In the following the results of the run on the same data of *TCA*, by using the indicator matrix and the Burt's table, respectively. As expected, the results are different.

```
> library(TCA)
> palIT=TCA(palI, Naxes=7, Graph=T);        # MTCA With applying TCA to the Indicator matrix
> palIT$VectMax
    Axe_1     Axe_2     Axe_3     Axe_4     Axe_5     Axe_6     Axe_7
0.5281647 0.4607483 0.4347083 0.4074407 0.3833644 0.2948703 0.1373551
>
> palIT$G
             Axe_1      Axe_2      Axe_3      Axe_4       Axe_5       Axe_6       Axe_7
2 Syl   -0.3113333 -0.2746721 -0.1564882 -0.1037305 -0.08484848 -0.10364365 -0.01486527
3 Syl    0.9686667  0.8498418  0.4963487  0.5180395  0.21688606  0.30988068 -0.51945198
```

```
4 Syl    0.9512035  0.8549878  0.4467209 -0.3311890  0.41556304  0.35844200  1.92275028
W Noun   0.2223799 -0.5586211  0.5066939 -0.2024872 -0.15426005  0.08159994 -0.02846000
W Verb  -0.9080971  0.5166481 -0.5773901  0.1950716  0.18280768  0.77542357  0.02909552
W Adj    0.5405185  0.7457273 -0.5467425  0.2610118  0.15807842 -1.12330526  0.03468563
T Child -0.7380000  0.1762183  0.4276635  0.8895575 -1.06898968 -0.20212999  0.22476288
T Revi  -0.5586452  0.1232286  0.6454322 -0.7830649  1.14645117 -0.28542125 -0.21147045
T Divu   0.5555968 -0.7484442 -0.5868348  0.7313428  0.57455876  0.13989026  0.10364560
T Summ   0.7328434  0.4608077 -0.4759592 -0.8563012 -0.65235275  0.34507925 -0.12035495
```

```
> palBT=TCA(Burt, Naxes=7, Graph=T);
> palBT$VectMax
    Axe_1      Axe_2      Axe_3      Axe_4      Axe_5      Axe_6      Axe_7
0.37105233 0.31613791 0.26083520 0.24865565 0.19230116 0.17480348 0.07949698

             Axe_1      Axe_2      Axe_3      Axe_4       Axe_5       Axe_6       Axe_7
2 Syl   -0.2326332 -0.1344975 -0.1296314 -0.1054135 -0.02477847 -0.02892970 -0.01433928
3 Syl    0.7072222  0.4191175  0.4189056  0.3321633  0.25224655  0.23263568 -0.23126395
4 Syl    0.7657788  0.4087712  0.3443630  0.3081762 -0.50555231 -0.38492626  0.95933607
W Noun   0.1839883 -0.3579042  0.2371689 -0.2028474  0.03557934  0.04154009  0.02058974
W Verb  -0.6816537  0.3361150 -0.2733233  0.2042174  0.21824299  0.25480609  0.12629706
W Adj    0.3641481  0.4716989 -0.2522624  0.2509870 -0.34689858 -0.40501585 -0.20074996
T Child -0.4683333  0.3793935  0.3298188 -0.3413051  0.25752073 -0.34968865 -0.04336199
T Revi  -0.3552366 -0.3214431  0.3464812  0.4533583 -0.41066625  0.17613085 -0.04371168
T Divu   0.3172846 -0.4125387 -0.4458319  0.2927228  0.40255031 -0.17265000  0.04284782
T Summ   0.5016426  0.3584008 -0.2232393 -0.4062870 -0.25855495  0.35109302  0.04353614
```

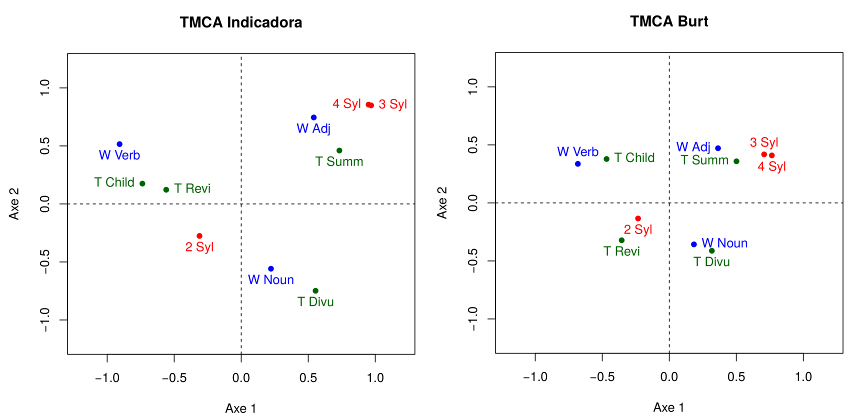

Figure 3: *The scatter plot of levels of Nardy (2007) palavras data, on the first factor plane issued by running the indicator matrix with TCA (Multiple Taxicab Correspondence Analysis) (left) and by running the Burt's matrix (right).*

In Figure 3 the patterns of the levels on the first factor plane issued by running the indicator matrix with TCA (Multiple Taxicab Correspondence Analysis) (left) and by running the Burt's matrix (right) are shown.

References

Abdi, H. (2007). Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition ($GSVD$). In: N. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage.

Ben Ammou, S., Saporta G. (1998). Sur la normalité asymptotique des valeurs propres en ACM sous l'hypothèse d'indépendance des variables. *Revue de Statistique Appliquée*, 46(3), 21-35.

Ben Ammou, S., Saporta G. (2003). On the connection between the distribution of eigenvalues in multiple correspondence analysis and log-linear models. *REVSTAT-Statistical Journal*, 1(0), 42-79.

Benzécri, J.P., et coll. (1973-82). *L'Analyse des données*, Tome 2. Paris: Dunod.

Benzécri, J.P. (1979). Sur les calcul des taux d'inertie dans l'analyse d'un questionnaire. *Les Cahiers de l'Analyse des Données*, 4(3), 377-379.

Camiz, S, Gomes, G.C. (2009). Correspondence Analyses for Studying the Language Complexity of Texts. VIII Congreso Chileno de Investigación Operativa, OPTIMA. Concepción (Chile), on CD-ROM.

Camiz, S, Gomes, G.C. (2013). Multiple and Joint Correspondence Analysis: Testing the True Dimension of a Study. *Modulad 2013, Modulad* 44: pp. 1-21.

Camiz, S., Gomes, G.C., Lemle, M., Nardy, M.N. (2009). Correspondence Analyses to Understand the Structure of Texts: a First Step towards a Structure Model. XLI Simposio Brasileiro de Pesquisa Operacional, Porto Seguro (Bahia, Brazil), on CD-ROM.

Choulakian, V., (2004). A Comparison of two Methods of Principal Component Analysis. In J. Antoch (Ed.): *Proceedings of Compstat2004 Symposium*. Berlin: Physica-Verlag/Springer, pp. 793-798.

Choulakian, V., (2006). Taxicab correspondence analysis. *Psychometrika*, 71(2): pp.333-345.

Choulakian, V., (2008). Multiple taxicab correspondence analysis. *Advances in Data Analysis and Classification*, 2(2): pp.177-206.

Eckart, C., Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211-218.

Greenacre, M.J. (1983). *Theory and Application of Correspondence Analysis*. London: Academic Press.

Greenacre, M.J. (1988). Correspondence analysis of mutlivariate categorical data by weighted least squares. *Biometrika*, 75, 457-467.

Greenacre, M.J. (2006). From Simple to Multiple Correspondence Analysis. In: Greenacre and Blasius (2006) (Eds.): pp. 41-76.

Greenacre, M.J., Blasius, J. (Eds.) (2006). *Multiple Correspondence Analysis and Related Methods*. Dordrecht (The Netherlands): Chapman and Hall (Kluwer).

Halle, M., Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In: Hale, K., Jay Keyser, S. (Eds.), *The View from Building*. Cambridge, MIT Press, 20: pp. 111-176.

Halle, M., Marantz, A. (1994). Some key features of Distributed Morphology. In: Carnie, A., Harley, H., (Eds.) *Papers on Phonology and Morphology*, Cambridge, MITWPL, 21: pp. 275-288.

Kendall, M.G., Stuart, A. (1961). *The Advanced Theory of Statistics*, vol. 2. London: Griffin.

Malinvaud, E. (1987). Data analysis in applied socio-economic statistics with special consideration of correspondence analysis. Marketing Science Conference, Joy en Josas: HEC-ISA.

Nardy, M.N.S. (2007). A sintaxe no interior das palavras - efeitos de gênero na lingua escrita contemporânea. PhD Thesis in Linguistics. Rio de Janeiro, Facultade de Leteras da Universidade Federal de Rio de Janeiro.

Nenadic, O., Greenacre, M. (2006). *Computation of multiple correspondence analysis, with code in R*. In: Greenacre and Blasius (2006) (Eds.), 523-551.

Nenadic, O., Greenacre, M. (2007). Correspondence analysis in R, with two- and three-dimensional graphics: the *ca* package. *Journal of Statistical Software*, 20(3), 1-13.

R-project (2009), http://www.r-project.org/

Orlóci, L. (1978). *Multivariate Analysis in Vegetation Research*, 2nd ed.. Den Haag: Junk.

Snee, R. D. (1974). "Graphical display of two-way contingency tables". *The American Statistician*, 28: pp. 912.