Multiple Correspondence Analysis in Marketing Research

Yangchun Du
Advisor: John C. Kern II
Department of Mathematics and Computer Science
Duquesne University
April 30, 2003
1. Introduction
2. Background
3. Details of Method
4. Simulated Data
5. MCA Properties
6. MSA Data
7. Conclusion and Future Work
8. References
9. S-Plus Code
Correspondence Analysis: A descriptive/exploratory technique to analyze simple two-way and multi-way tables containing some measure of correspondence between the rows and columns.

Goal: Convert the numerical information from a contingency table into a two-dimensional graphical display.

Data Type: Categorical Data.

Application Area: Marketing Research.

Advantage: Allowing researcher to visualize relationships among categories of categorical variables for large data sets.
Multiple Correspondence Analysis (MCA) is considered to be an extension of simple correspondence analysis to more than $Q = 2$ variables.

**Indicator matrix $Z$ ($n \times \sum_q J_q$)**

Row—individuals (usually people).
Column—category of a categorical variable.

$$Z = \begin{bmatrix}
    \text{male} & \text{female} & \text{location1} & \text{location2} \\
    1 & 0 & 1 & 0 \\
    1 & 0 & 0 & 1 \\
    0 & 1 & 1 & 0 \\
    0 & 1 & 0 & 1 \\
    0 & 1 & 0 & 1
\end{bmatrix}$$

**Burt Matrix $B = Z^t \times Z$**

$$B = \begin{bmatrix}
    \text{male} & \text{female} & \text{location1} & \text{location2} \\
    \text{male} & 2 & 0 & 1 & 1 \\
    \text{female} & 0 & 3 & 1 & 2 \\
    \text{location1} & 1 & 1 & 2 & 0 \\
    \text{location2} & 1 & 2 & 0 & 3
\end{bmatrix}$$
**Background**

**Computation for Simple Correspondence Analysis**

Let $\mathbf{N}$ be a $I \times J$ matrix representing a contingency table of two categorical variables.

- **Row mass $r_i$**: row sums divided by grand total $n$, $r_i = \frac{n_i}{n}$; vector of row masses $r$.

- **Column mass $c_j$**: column sums divided by grand total $n$, $c_j = \frac{n_j}{n}$; vector of column masses $c$.

- **Correspondence Matrix $\mathbf{P}$**: original table $\mathbf{N}$ divided by grand total $n$, $\mathbf{P} = \frac{\mathbf{N}}{n}$.

- **Row profiles**: rows of the original table $\mathbf{N}$ divided by respective row totals; equivalently $\mathbf{D}_r^{-1}\mathbf{P}$, where $\mathbf{D}_r$ is diagonal matrix of row masses.

- **Column profiles**: columns of the original table $\mathbf{N}$ divided by respective column totals; equivalently $\mathbf{PD}_c^{-1}$, where $\mathbf{D}_c$ is diagonal matrix of column masses.

- **Standardized Residuals**: $I \times J$ matrix $\mathbf{A}$ with elements $a_{ij} = \frac{p_{ij} - r_i c_j}{\sqrt{r_i c_j}}$; $\mathbf{A} = \mathbf{D}_r^{-1/2}(\mathbf{P} - r c^T)\mathbf{D}_c^{-1/2}$. 
• **Singular Value Decomposition** (SVD) of $I \times J$ matrix $A$ into product of three matrices:

$$A = U \Gamma V^T,$$

$\Gamma$ diagonal matrix $\gamma_{11} \geq \gamma_{22} \geq \cdots \geq \gamma_{kk} > 0$ (these are singular values); columns of matrices $U$ and $V$ are left and right singular vectors, respectively.

• **Chi-square statistic**:

$$\chi^2 = n \sum_i \sum_j a_{ij}^2.$$

• **Total Inertia**:

$$\sum_i \sum_j a_{ij}^2.$$ Equivalently $\frac{\chi^2}{n}$.

• Maximum $K$ dimensions for graphical display in CA, where $K = \min\{I-1, J-1\}$. Squares of singular values of $A$ also decompose total inertia: $\lambda_1 \ldots \lambda_k$, are **principal inertias**. 

Greenacre (1984) shows that the correspondence analysis of the indicator matrix $Z$ are identical to those in the analysis of $B$. Furthermore, the principal inertias of $B$ are squares of those of $Z$.

• The **principal coordinates of the rows** are obtained as $D_r^{-1/2} U \Gamma$.

• The **principal coordinates of the columns** are obtained as $D_c^{-1/2} V \Gamma$. 
Details of Method

• Currently, some statistical software packages can perform MCA.
  
  – *SAS*: built-in *corresp* procedure
  
  – *SPSS Categories*: CA procedure

  **Commonality**: Decompose Burt matrix using SVD.

• Greenacre (1988)
  
  – Creates modified Burt matrix— the original Burt matrix with modified sub-matrices on its diagonal
  
  – Advantage (over standard MCA analysis): greater percentage of explained variation (total inertia) by two-dimensional solution for some categorical datasets.
weighted least-squares approximation of a Burt matrix,

\[ B \approx nrr^T + nDXD_\beta X^T D \]

where \( n \) is the grand total, \( r \) is the row mass, \( D \) is the diagonal matrix of the mass. Let

\[ S = n^{-1/2}D^{-1/2}BD^{-1/2} \]

so the SVD of \( S \) is

\[ S = UD_\alpha V^T \]

\[
\begin{bmatrix}
1 & 0 \\
0 & D_\beta
\end{bmatrix}
= n^{-1/2}D_\alpha
\]

(1)

\[
\begin{bmatrix}
1 & X
\end{bmatrix}
= D^{-1/2}U
\]

(2)

We then use \( D_\beta \) and \( X \) to obtain \( \sum_q J_q \times K \) matrix \( \Xi \) of coordinates:

\[ \Xi = XD_\beta^{1/2} \]

Columns 1 and 2 of \( \Xi \) are the category-representing coordinates.
We build a model for the whole matrix $B - nrr^T$, namely
\[
B - nrr^T \approx nD\hat{X}_\beta X^T D + C
\]
where $C$ is a block diagonal matrix with sub-matrices $C_{qq}$ ($q = 1, \ldots, Q$) down the diagonal and zeros elsewhere. So, the new sub-matrix $N^*_{qq}$ on the diagonal of $B^*$, which is given by
\[
N^*_{qq} = nr_q r^T_q + nD_q X_q D_\beta X^T D_q,
\] 
has the same row and column margins as $N_{qq}$, where the vector of $J_q$ masses for variable $q$ is denoted by $r_q$. The $J_q \times J_q$ diagonal matrix formed from the elements of $r_q$ is now denoted by $D_q$. Meanwhile, the diagonal matrix $D_\beta$ contains a scale parameter for each dimension. The parameter $X$ is partitioned two-wise according to the variable as $X_1 \cdots X_Q$. so $X_q$ is a $J_q \times K$ sub matrices.

The procedure of this algorithm:

1. Start with a solution for $X$ and $D_\beta$ based on MCA.
2. Replace the sub matrices on the diagonal of $B$ with those “estimated” by $X$ and $D_\beta$ given by (3).
3. Perform a correspondence analysis on the modified matrix $B^*$, setting $X$ equal to the first $K$ vectors of optimal row or column parameters and the diagonal
of $D_\beta$ equal to the square roots of the first $K$ principal inertias respectively.

4. Go back to 2 and repeat until the iterations converge; that is, when the decrease in the discrepancy function from iteration to iteration is practically zero.
### Simulated Data

#### Table 1: Simulated Burt Matrix

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>nonHS</th>
<th>HS</th>
<th>College</th>
<th>locA</th>
<th>locB</th>
<th>locC</th>
<th>BrandX</th>
<th>BrandY</th>
<th>BrandZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>496</td>
<td>0</td>
<td>197</td>
<td>259</td>
<td>40</td>
<td>371</td>
<td>72</td>
<td>53</td>
<td>303</td>
<td>96</td>
<td>97</td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
<td>504</td>
<td>39</td>
<td>216</td>
<td>249</td>
<td>310</td>
<td>98</td>
<td>96</td>
<td>256</td>
<td>129</td>
<td>131</td>
</tr>
<tr>
<td>nonHS</td>
<td>197</td>
<td>39</td>
<td>236</td>
<td>0</td>
<td>0</td>
<td>212</td>
<td>12</td>
<td>12</td>
<td>171</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>HS</td>
<td>259</td>
<td>216</td>
<td>0</td>
<td>475</td>
<td>0</td>
<td>330</td>
<td>77</td>
<td>68</td>
<td>257</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>College</td>
<td>40</td>
<td>249</td>
<td>0</td>
<td>0</td>
<td>289</td>
<td>139</td>
<td>81</td>
<td>69</td>
<td>131</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>locA</td>
<td>371</td>
<td>310</td>
<td>212</td>
<td>330</td>
<td>139</td>
<td>681</td>
<td>0</td>
<td>0</td>
<td>515</td>
<td>82</td>
<td>82</td>
</tr>
<tr>
<td>locB</td>
<td>72</td>
<td>98</td>
<td>12</td>
<td>77</td>
<td>81</td>
<td>0</td>
<td>170</td>
<td>0</td>
<td>22</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>locC</td>
<td>53</td>
<td>96</td>
<td>12</td>
<td>68</td>
<td>69</td>
<td>0</td>
<td>0</td>
<td>149</td>
<td>22</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>BrandX</td>
<td>303</td>
<td>256</td>
<td>171</td>
<td>257</td>
<td>131</td>
<td>515</td>
<td>22</td>
<td>22</td>
<td>559</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BrandY</td>
<td>96</td>
<td>125</td>
<td>34</td>
<td>102</td>
<td>85</td>
<td>82</td>
<td>126</td>
<td>13</td>
<td>0</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>BrandZ</td>
<td>97</td>
<td>123</td>
<td>31</td>
<td>116</td>
<td>73</td>
<td>84</td>
<td>22</td>
<td>114</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

#### Table 2: Simulated Modified Burt Matrix

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>nonHS</th>
<th>HS</th>
<th>College</th>
<th>locA</th>
<th>locB</th>
<th>locC</th>
<th>BrandX</th>
<th>BrandY</th>
<th>BrandZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>272</td>
<td>234</td>
<td>197</td>
<td>259</td>
<td>40</td>
<td>371</td>
<td>72</td>
<td>53</td>
<td>303</td>
<td>96</td>
<td>97</td>
</tr>
<tr>
<td>Female</td>
<td>234</td>
<td>289</td>
<td>39</td>
<td>216</td>
<td>249</td>
<td>310</td>
<td>98</td>
<td>96</td>
<td>256</td>
<td>129</td>
<td>131</td>
</tr>
<tr>
<td>nonHS</td>
<td>197</td>
<td>39</td>
<td>68</td>
<td>114</td>
<td>59</td>
<td>212</td>
<td>12</td>
<td>12</td>
<td>171</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>HS</td>
<td>259</td>
<td>216</td>
<td>114</td>
<td>238</td>
<td>134</td>
<td>330</td>
<td>77</td>
<td>68</td>
<td>257</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>College</td>
<td>40</td>
<td>249</td>
<td>59</td>
<td>134</td>
<td>101</td>
<td>139</td>
<td>81</td>
<td>69</td>
<td>131</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>locA</td>
<td>371</td>
<td>310</td>
<td>212</td>
<td>330</td>
<td>139</td>
<td>345</td>
<td>117</td>
<td>115</td>
<td>515</td>
<td>82</td>
<td>82</td>
</tr>
<tr>
<td>locB</td>
<td>72</td>
<td>98</td>
<td>12</td>
<td>77</td>
<td>81</td>
<td>117</td>
<td>67</td>
<td>45</td>
<td>22</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>locC</td>
<td>53</td>
<td>96</td>
<td>12</td>
<td>68</td>
<td>69</td>
<td>115</td>
<td>45</td>
<td>66</td>
<td>22</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>BrandX</td>
<td>303</td>
<td>256</td>
<td>171</td>
<td>257</td>
<td>131</td>
<td>515</td>
<td>22</td>
<td>22</td>
<td>502</td>
<td>98</td>
<td>98</td>
</tr>
<tr>
<td>BrandY</td>
<td>96</td>
<td>125</td>
<td>34</td>
<td>102</td>
<td>85</td>
<td>82</td>
<td>126</td>
<td>13</td>
<td>98</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>BrandZ</td>
<td>97</td>
<td>123</td>
<td>31</td>
<td>116</td>
<td>73</td>
<td>84</td>
<td>22</td>
<td>114</td>
<td>87</td>
<td>21</td>
<td>21</td>
</tr>
</tbody>
</table>
### Table 3: Summary for simulated data

<table>
<thead>
<tr>
<th>k</th>
<th>Burt Matrix Principal inertia</th>
<th>Burt Matrix Percent</th>
<th>Modified Burt Matrix Principal inertia</th>
<th>Modified Burt Matrix Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4843487</td>
<td>27.67707</td>
<td>0.2712651</td>
<td>29.59681</td>
</tr>
<tr>
<td>2</td>
<td>0.3874350</td>
<td>22.13915</td>
<td>0.1691768</td>
<td>18.4583</td>
</tr>
<tr>
<td>3</td>
<td>0.2988519</td>
<td>17.07725</td>
<td>0.1278038</td>
<td>13.94424</td>
</tr>
<tr>
<td>4</td>
<td>0.2463057</td>
<td>14.07461</td>
<td>0.1044092</td>
<td>11.39173</td>
</tr>
<tr>
<td>5</td>
<td>0.1253908</td>
<td>7.165191</td>
<td>0.09827268</td>
<td>10.7222</td>
</tr>
<tr>
<td>6</td>
<td>0.0112663</td>
<td>6.437910</td>
<td>0.09176939</td>
<td>10.01265</td>
</tr>
<tr>
<td>7</td>
<td>0.0095004</td>
<td>5.428824</td>
<td>0.02237811</td>
<td>2.4416</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0.02048708</td>
<td>2.235275</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0.006054817</td>
<td>0.6606204</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0.004917945</td>
<td>0.5365802</td>
</tr>
</tbody>
</table>
Simulated Data-Continued

MCA Graphical display of Burt Matrix for Simulated Data

MCA Graphical display of Modified Burt Matrix for simulated data
Lemma: Burt Matrix of duplicated data is 2 times of that of the original data.

Theorem: The MCA for Burt Matrix $B$ is identical to MCA for Burt matrix $B^* = k \cdot B$ for any $k > 0$.

Proof:

Let $Z$ is a $m \times n$ indicator matrix, binary representing the data with $n$ categorical variables and $m$ cases (observations).

$$Z = \begin{bmatrix}
Z_{11} & Z_{12} & \cdots & Z_{1n} \\
Z_{21} & Z_{22} & \cdots & Z_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
Z_{m1} & Z_{m2} & \cdots & Z_{mn}
\end{bmatrix}$$

the transpose of $Z$

$$Z^T = \begin{bmatrix}
Z_{11} & Z_{21} & \cdots & Z_{m1} \\
Z_{21} & Z_{22} & \cdots & Z_{m2} \\
\vdots & \vdots & \ddots & \vdots \\
Z_{1n} & Z_{2n} & \cdots & Z_{mn}
\end{bmatrix}$$
\[ n \times n \text{ Burt Matrix } B = z^T \times z = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nn} \end{bmatrix} \]

where

\[ B_{11} = Z_{11}Z_{11} + Z_{12}Z_{12} + \cdots + Z_{1n}Z_{1n} \]
\[ B_{12} = Z_{11}Z_{12} + Z_{21}Z_{22} + \cdots + Z_{m1}Z_{1n} \]
\[ \vdots \]
\[ B_{nn} = Z_{1n}Z_{1n} + Z_{2n}Z_{2n} + \cdots + Z_{mn}Z_{mn} \]

If we duplicate the data, that means \( Z^* \) is \( 2m \times n \) indicator matrix as follows.

\[
Z^* = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{m1} & Z_{m2} & \cdots & Z_{mn} \\ Z_{11} & Z_{12} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{m1} & Z_{m2} & \cdots & Z_{mn} \end{bmatrix}
\]

\( B^* \) is still a \( n \times n \) symmetric matrix

\[
B^* = Z^{*T} \times Z^* = \begin{bmatrix} 2B_{11} & 2B_{12} & \cdots & 2B_{1n} \\ 2B_{21} & 2B_{22} & \cdots & 2B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 2B_{n1} & 2B_{n2} & \cdots & 2B_{nn} \end{bmatrix} = 2 \times B
\]
Lemma and theorem—Continue

**Theorem** : The MCA for Burt Matrix $\mathbf{B}$ is identical to MCA for Burt matrix $\mathbf{B}^* = k \cdot \mathbf{B}$ for any $k > 0$.

**Proof**:

Let $\mathbf{B}$ is an $I \times I$ Burt matrix, with row and column total $B_{i+} (i = 1, \ldots, I)$ and grand total $n$. Let $r$ be vectors of elements $r_i = \frac{B_{i+}}{n}$, call row masses. $\mathbf{D}$ is the diagonal metrics of these masses. As we state early in MCA, the expected frequencies $e_{ii} = \frac{B_{i+}B_{+i}}{n} = nr_ir_i$. We also let $\mathbf{S} = n^{-1/2}\mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}$ with $s_{ii} = \frac{B_{ii}}{\sqrt{e_{ii}}}$.

So the SVD of $\mathbf{S}$ is $\mathbf{S} = \mathbf{U} \mathbf{D}_\alpha \mathbf{V}^T$

\[
\begin{bmatrix}
1 & 0 \\
0 & \mathbf{D}_\beta
\end{bmatrix} = n^{-1/2} \mathbf{D}_\alpha \tag{4}
\]

\[
\begin{bmatrix}
1 & \mathbf{X}
\end{bmatrix} = \mathbf{D}^{-1/2} \mathbf{U} \tag{5}
\]

Actually, in MCA, the element in the diagonal of $\mathbf{D}_\beta$ is the value of principal inertia, the $\mathbf{X}\sqrt{\mathbf{D}_\beta}$ is the value of principal coordinate. In this theorem, we need to prove
that $D_\beta$ and $X$ of $B^*$ are the same as that of $B$.

For $B^*$, row mass $r^*_i = \frac{B^*_i}{n^*} = \frac{kB_i}{kn} = r_i$ So, $e^*_{ii} = kn r_i r_i = ke_{ii}$.

for each element $S^*_{ii}$ in $S^*$, $S^*_{ii} = \frac{B^*_{ii}}{\sqrt{e^*_{ii}}} = \frac{kB_{ii}}{\sqrt{ke_{ii}}} = \sqrt{k}S_{ii}$

once again, the SVD of $S$ is $S^* = U^* D^*_\alpha V^*T$, we also know

$U^* = U$, $D^*_\alpha = \sqrt{k}D_\alpha$, and $V^*T = V$.

As we can see, in the MCA of $B^*$, the right of equation [2] keeps the same since $(kn)^{-1/2}\sqrt{k}D_\alpha = n^{-1/2}D_\alpha$, and the right side of equation [3]($D^{-1/2}U$) also keeps the same.

$D^*_\beta = D_\beta$, $X^* = X$. 
### MSA Data

<table>
<thead>
<tr>
<th>Brand1</th>
<th>Brand2</th>
<th>Male</th>
<th>Female</th>
<th>College</th>
<th>High School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand1</td>
<td>695</td>
<td>0</td>
<td>373</td>
<td>322</td>
<td>499</td>
</tr>
<tr>
<td>Brand2</td>
<td>0</td>
<td>842</td>
<td>456</td>
<td>386</td>
<td>420</td>
</tr>
<tr>
<td>Male</td>
<td>373</td>
<td>456</td>
<td>829</td>
<td>0</td>
<td>457</td>
</tr>
<tr>
<td>Female</td>
<td>322</td>
<td>386</td>
<td>0</td>
<td>708</td>
<td>462</td>
</tr>
<tr>
<td>College</td>
<td>499</td>
<td>420</td>
<td>457</td>
<td>462</td>
<td>919</td>
</tr>
<tr>
<td>High School</td>
<td>196</td>
<td>422</td>
<td>372</td>
<td>246</td>
<td>335</td>
</tr>
</tbody>
</table>

### Table 4: Burt Matrix

<table>
<thead>
<tr>
<th>Brand1</th>
<th>Brand2</th>
<th>Male</th>
<th>Female</th>
<th>College</th>
<th>High School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand1</td>
<td>346</td>
<td>349</td>
<td>373</td>
<td>322</td>
<td>499</td>
</tr>
<tr>
<td>Brand2</td>
<td>349</td>
<td>493</td>
<td>456</td>
<td>386</td>
<td>420</td>
</tr>
<tr>
<td>Male</td>
<td>373</td>
<td>456</td>
<td>468</td>
<td>361</td>
<td>457</td>
</tr>
<tr>
<td>Female</td>
<td>322</td>
<td>386</td>
<td>361</td>
<td>347</td>
<td>462</td>
</tr>
<tr>
<td>College</td>
<td>499</td>
<td>420</td>
<td>457</td>
<td>462</td>
<td>584</td>
</tr>
<tr>
<td>High School</td>
<td>196</td>
<td>422</td>
<td>372</td>
<td>246</td>
<td>335</td>
</tr>
</tbody>
</table>

### Table 5: Modified Burt Matrix

<table>
<thead>
<tr>
<th>Singular Value</th>
<th>Principal Inertia</th>
<th>Chi-Square</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burt k=1</td>
<td>0.64473</td>
<td>0.41568</td>
<td>1993.5</td>
<td>41.57</td>
</tr>
<tr>
<td>matrix k=2</td>
<td>0.57628</td>
<td>0.33210</td>
<td>1592.67</td>
<td>33.21</td>
</tr>
<tr>
<td>k=3</td>
<td>0.50222</td>
<td>0.25233</td>
<td>1209.64</td>
<td>25.22</td>
</tr>
<tr>
<td>total</td>
<td>1</td>
<td>4795.81</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Modified k=1</td>
<td>0.33318</td>
<td>0.11101</td>
<td>182.735</td>
<td>85.6</td>
</tr>
<tr>
<td>Burt k=2</td>
<td>0.13667</td>
<td>0.01868</td>
<td>30.749</td>
<td>14.4</td>
</tr>
<tr>
<td>Matrix total</td>
<td>0.12969</td>
<td>213.484</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6: Summary for MCA of Burt and Modified Burt Matrix
MSA Data-Continued

MCA Graphical display of Burt Matrix for MSA Data

- First Principal Axis 41.57%
- Second Principal Axis 33.21%

MCA Graphical display of Modified Burt Matrix for MSA Data

- First Principal Axis 85.6%
- Second Principal Axis 14.4%
Conclusion and Future Work

Identify the characteristics of data that work well with Greenacre method.


Function to calculate the Square Root of Matrix A

\[
pdsroot \leftarrow \text{function}(A) \\
\{ \\
p \leftarrow \text{dim}(A)[1] \quad \# \text{dimension of matrix} \\
eigenv \leftarrow \text{eigen}(A, \text{symmetric} = T) \quad \# \text{spectral decomposition of } A \\
\text{rootA} \leftarrow \text{eigenv}$\text{vectors} \times \text{sqrt(eigenv}$\text{values})) \times \text{t(eigenv}$\text{vectors}) \\
\text{rootA} \\
\}
\]

Perform MCA for Burt matrix in S-plus

\[
\text{r_apply}(N,1,\text{sum}) \quad \# \text{N is the i by i Burt Matrix} \\
\text{n_sum}(r) \\
\text{r_apply}(N,1,\text{sum})/n \\
\text{c_apply}(N,2,\text{sum})/n \\
\text{Dr}_\text{diag}(r) \\
\text{Dc}_\text{diag}(c) \\
S_\left((1/\text{sqrt}(n))\times\text{solve}(pdsroot(Dr))\times\text{solve}(pdsroot(Dc)) \right) \\
\text{U_svd}(S)$u \\
\text{V_svd}(S)$v \\
\text{Dalph} \leftarrow \text{diag(svd(S)$d)} \\
\text{Dmu}_\left(\text{sqrt}(1/n)\times\text{Dalph}[2:(i-1),2:(i-1)]) \quad \# \text{Dmu is the principal inertia} \\
\text{X_pdsroot}(\text{Dr}) \\
\text{X_solve}(X)$u \\
\text{X}_\times X[1,1] \\
\text{X}_\times X[1:(i-1),2:(i-1)] \\
\text{Y_pdsroot}(\text{Dc}) \\
\text{Y_solve}(Y)$v \\
\text{Y}_\times Y[1,1] \\
\text{Y}_\times Y[1:(i-1),2:(i-1)] \\
d\text{yc}_\times n\times(r\times t(c)+\text{Dr}\times\text{X}\times\text{sqrt(Dmu)}\times t(Y)\times\text{Dc}) \quad \# \text{dyc is the Weight} \\
\quad \# \text{least-squares approximation of Burt matrix} \\
\text{fs} \leftarrow X \times \text{sqrt(Dmu)} \quad \# \text{this is the principal coordinate} \\
\quad \# \text{for row and column (note: row and column is same in this case)}
Iterative Algorithm to obtain the Modified Burt Matrix for MSA Data (number of observation is 1537), B is 6 by 6 Burt Matrix, and have three categorical variables, each has 2 variables.

```r
N<-B
for (i in 1:5)
{
  r_apply(N,1,sum)
  n_sum(r)
  r_apply(N,1,sum)/n
  c_apply(N,2,sum)/n
  Dr_diag(r)
  Dc_diag(c)
  S_(1/sqrt(n))*solve(pdsroot(Dr))%*%N%*%solve(pdsroot(Dc))
  U_svd(S)$u
  V_svd(S)$v
  Dalph<-diag(svd(S)$d)
  Dmu_(sqrt(1/n)*Dalph)[2:6,2:6]
  X_pdsroot(Dr)
  X_solve(X)$U
  X_X/X[1,1]
  X_X[1:6,2:6]
  Y_pdsroot(Dc)
  Y_solve(Y)$V
  Y_Y/Y[1,1]
  Y_Y[1:6,2:6]
  dyc_n*(r%*%t(c)+Dr%*%X%*%Dmu%*%t(Y))%*%Dc
}
Dbeta<-sqrt(Dmu)
N<-B

N11<-N[1:2,1:2]
R11<-apply(N11,1,sum)
D11<-diag(R11)
X11<-X[1:2,1:5]
N11star<-round(R11%*%t(R11)/1537)
N11star<-N11star+round(D11%*%X11%*%Dbeta%*%t(X11))%*%D11/n

R22<-apply(N22,1,sum)
D22<-diag(R22)
```
X22 <- X[3:4, 1:5]
N22star <- round(R22%*%t(R22)/1537)
N22star <- N22star + round(D22%*%X22%*%Dbeta%*%t(X22)%*%D22/n)

R33 <- apply(N33, 1, sum)
D33 <- diag(R33)
X33 <- X[5:6, 1:5]
N33star <- round(R33%*%t(R33)/1537)
N33star <- N33star + round(D33%*%X33%*%Dbeta%*%t(X33)%*%D33/n)

N[1:2, 1:2] <- N11star

Plot two-dimensional graphical display

fsname <- c('Brand1', 'Brand2', 'Male', 'Female', 'College', 'High School')

corrplot <- function(fs, fsname)
{
  xlabes <- fsname
  plot(fs[, 1], fs[, 2], pch = '*', xlab = paste("First Principal Axis"),
       ylab = paste("Second Principal Axis"))
  text(dycfs[, 1] - 0.01, dycfs[, 2] - 0.03, labels = xlabes, adj = 0)
  title(main = "MCA Graphical display of Modified Burt Matrix for MSA Data")
  abline(h = 0, v = 0)
  return(fs = fs[, c(1, 2)])
}