

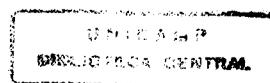
UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA  
DEPARTAMENTO DE ESTATÍSTICA

**ANÁLISE DE CORRESPONDÊNCIA  
PARA DADOS COM  
ESTRUTURA DE GRUPO**

**ADMUR SEVERINO PAMPLONA**

Campinas - SP

1998



**ANÁLISE DE CORRESPONDÊNCIA  
PARA DADOS COM  
ESTRUTURA DE GRUPO**

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendido por ADMUR SEVERINO PAMPLONA e aprovado pela Comissão Julgadora.

Campinas, 12 de fevereiro de 1998



---

Prafa. Dra. REGINA CÉLIA C. P. MORAN  
Orientadora

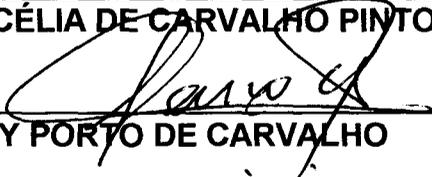
Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas, como requisito parcial para obtenção do Título de MESTRE EM ESTATÍSTICA.

Dissertação de Mestrado defendida e aprovada em 12 de fevereiro de 1998  
pela Banca Examinadora composta pelos Profs. Drs.



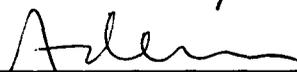
---

Prof (a). Dr (a). REGINA CÉLIA DE CARVALHO PINTO MORAN



---

Prof (a). Dr (a). JOSÉ RUY PORTO DE CARVALHO



---

Prof (a). Dr (a). ADEMIR JOSÉ PETENATE

Dedico aos  
meus três amores:  
Wanderleya Nara,  
Ana Clara e  
Lucas Henrique.

Agradeço:

à Regina, pela orientação, compreensão e incentivo;

aos membros da banca examinadora, José Ruy e o Prof. Ademir Petenate, pela correção e sugestões.

à Clarice, Carla e Cristiano, pelo apoio nos momentos críticos;

à Wanderleya, pelas críticas, sugestões e digitação;

à Coleta, pelos dados da aplicação e pelas sugestões;

aos meus pais, José e Elzira, pelo amor e conhecimento que me transmitiram e transmitem;

aos colegas da Estatística, pelos momentos compartilhados;

aos colegas do Departamento de Matemática do ICLMA/UFMT, pelo apoio silencioso e por acreditarem na realização deste trabalho;

à Irmã Maria José pela correção ortográfica e carinho e ao Pedro Paulo pela amizade;

a todos que de alguma forma contribuíram para a execução deste trabalho.

“... a graphical description is more easily assimilated and interpreted than a numerical one and can assist all three junctions mentioned

- summarizing a large mass of numerical data
- simplifying the aspect of the data by appealing our natural ability to absorb visual images
- and (hopefully) providing a global view of the information, thereby stimulating possible explanations

... rising importance of exploratory data analysis in a world which grows more complex and varied each day...”

Greenacre (1984)

## SUMÁRIO

### 1. INTRODUÇÃO

|   |   |
|---|---|
| <b>1.1 APRESENTAÇÃO</b> .....               | 1 |
| <b>1.2 ANÁLISE DE CORRESPONDÊNCIA</b> ..... | 4 |
| 1.2.1 A TÉCNICA .....                       | 4 |
| 1.2.2 APLICAÇÃO .....                       | 5 |
| 1.2.3 UM BREVE HISTÓRICO .....              | 7 |

### 2. ANÁLISE DE CORRESPONDÊNCIA SIMPLES

|   |    |
|---|----|
| <b>2.1 INTRODUÇÃO</b> .....                       | 8  |
| <b>2.2 OS DADOS</b> .....                         | 10 |
| <b>2.3 RELAÇÕES ENTRE AS LINHAS</b> .....         | 13 |
| 2.3.1 A NUVEM DE PONTOS NO ESPAÇO $R^J$ .....     | 13 |
| 2.3.2 A MÉTRICA DA NUVEM NO ESPAÇO $R^J$ .....    | 16 |
| 2.3.2.1 A INÉRCIA .....                           | 17 |
| 2.3.2.2 O ESPAÇO TRANSFORMADO .....               | 20 |
| 2.3.3 O AJUSTE DO SUBESPAÇO .....                 | 22 |
| 2.3.3.1 ESPAÇO VETORIAL NÃO CENTRALIZADO .....    | 26 |
| <b>2.4 RELAÇÕES ENTRE AS COLUNAS</b> .....        | 29 |
| 2.4.1 A NUVEM DE PONTOS NO ESPAÇO $R^I$ .....     | 29 |
| 2.4.2 A MÉTRICA DA NUVEM NO ESPAÇO $R^I$ .....    | 30 |
| 2.4.2.1 INÉRCIA TOTAL E ESPAÇO TRANSFORMADO ..... | 31 |
| 2.4.3 O AJUSTE DO SUBESPAÇO .....                 | 33 |
| 2.4.3.1 ESPAÇO VETORIAL NÃO CENTRALIZADO .....    | 33 |

|  |           |
|--|-----------|
| <b>2.5 RELAÇÕES ENTRE LINHAS E COLUNAS .....</b>   | <b>35</b> |
| 2.5.1 MASSAS E CENTRÓIDES.....   | 35        |
| 2.5.2 ESPAÇOS VETORIAIS.....   | 36        |
| 2.5.3 FÓRMULAS DE TRANSIÇÃO .....  | 37        |
| <b>2.6 INTERPRETAÇÃO.....</b>  | <b>39</b> |
| 2.6.1 INTERPRETAÇÃO DOS EIXOS .....  | 39        |
| 2.6.2 INTERPRETAÇÃO DOS PONTOS NO CONJUNTO DOS EIXOS .....   | 43        |
| 2.6.3 PONTOS SUPLEMENTARES.....  | 46        |
| <br>   |           |
| <b>3. ANÁLISE DE CORRESPONDÊNCIA EM GRUPOS</b>   |           |
| <br>   |           |
| <b>3.1 INTRODUÇÃO .....</b>  | <b>47</b> |
| <b>3.2 OS DADOS .....</b>  | <b>49</b> |
| <b>3.3 RELAÇÕES ENTRE AS MATRIZES <math>N</math>, <math>N_1</math>, <math>N_2</math> E <math>N_3</math>.....</b> | <b>53</b> |
| 3.3.1 AS NUVENS DE PONTOS E SUAS RELAÇÕES .....  | 53        |
| 3.3.2 AS MÉTRICAS E SUAS RELAÇÕES .....  | 57        |
| 3.3.2.1 <i>RELAÇÕES ENTRE AS INÉRCIAS TOTAIS DAS NUVENS</i> .....  | 59        |
| 3.3.3 OS AJUSTES DAS NUVENS E SUAS RELAÇÕES.....   | 61        |
| 3.3.3.1 <i>RELAÇÕES ENTRE AS INÉRCIAS DOS EIXOS</i> .....  | 63        |
| <br>   |           |
| <b>4. APLICAÇÃO DA ANÁLISE DE CORRESPONDÊNCIA EM GRUPOS</b>  |           |
| <br>   |           |
| <b>4.1 INTRODUÇÃO.....</b>   | <b>66</b> |
| <b>4.2 OS DADOS .....</b>  | <b>67</b> |
| <b>4.3 TRANSFORMAÇÃO DOS DADOS .....</b>   | <b>74</b> |
| 4.3.1 APLICAÇÃO DO ALGORITMO DA AC .....   | 74        |
| <b>4.4 DESCRIÇÃO.....</b>  | <b>78</b> |
| 4.4.1 DESCRIÇÃO DO PRIMEIRO EIXO.....  | 78        |

|                                |   |            |
|--------------------------------|---|------------|
| 4.4.2                          | DESCRIÇÃO DO SEGUNDO EIXO.....                              | 81         |
| 4.4.3                          | DESCRIÇÃO DO GRÁFICO DOS CENTRÓIDES.....                    | 82         |
| 4.4.4                          | DESCRIÇÃO DOS PONTOS SUPLEMENTARES.....                     | 85         |
| <br>                           |   |            |
| <b>5. CONSIDERAÇÕES FINAIS</b> |   |            |
| 5.1                            | <i>SOBRE O TRABALHO.....</i>                                | 91         |
| 5.2                            | <i>SOBRE A TÉCNICA .....</i>                                | 92         |
| <br>                           |   |            |
| <b>APÊNDICE A</b>              |   |            |
| A.1                            | <i>ACS DA MATRIZ N.....</i>                                 | 94         |
| A.2                            | <i>ACS DA MATRIZ N<sub>1</sub>.....</i>                     | 97         |
| A.3                            | <i>ACS DA MATRIZ N<sub>2</sub>.....</i>                     | 100        |
| A.4                            | <i>ACS DA MATRIZ N<sub>3</sub>.....</i>                     | 103        |
| <br>                           |   |            |
| <b>APÊNDICE B</b>              |   |            |
| B.1                            | <i>DESCRIÇÃO PARA O ALGORITMO DA AC .....</i>               | 107        |
| B.2                            | <i>PROGRAMA NO IML/SAS DO ALGORITMO DA AC .....</i>         | 123        |
| B.3                            | <i>SAÍDA DO PROGRAMA NO IML/SAS DO ALGORITMO DA AC.....</i> | 130        |
| B.4                            | <i>PROGRAMA DE AC COM O PROC CORRESPO/SAS .....</i>         | 137        |
| B.5                            | <i>SAÍDA DO PROGRAMA DE AC COM O PROC CORRESPO/SAS.....</i> | 139        |
| B.6                            | <i>PROGRAMA DE AC DO SAS PARA PONTOS SUPLEMENTARES.....</i> | 145        |
| B.7                            | <i>SAÍDA DO PROGRAMA/SAS PARA PONTOS SUPLEMENTARES.....</i> | 148        |
| <br>                           |   |            |
| <b>BIBLIOGRAFIA.....</b>       |   | <b>159</b> |

## LISTA DE FIGURAS

|  |    |
|--|----|
| FIGURA 1: Representação do perfil linhas $R_i$ no espaço $\mathbf{R}^3$ .....        | 14 |
| FIGURA 2: Representação do perfil linhas transformado no espaço $\mathbf{R}^3$ ..... | 21 |
| FIGURA 3: Decomposição da distância $d(R_i, c)$ .....                                | 40 |
| FIGURA 4: Representação esquemática das matrizes $N$ , $N_1$ , $N_2$ e $N_3$ .....   | 52 |

## LISTA DE TABELAS

|   |    |
|---|----|
| TABELA 1: Modelo de Tabela de Contingência Simples .....                                  | 10 |
| TABELA 2: Decomposição da Inércia dos Perfis Linhas sobre os Eixos Principais .....       | 41 |
| TABELA 3: Decomposição da Inércia dos Perfis Colunas sobre os Eixos Principais .....      | 41 |
| TABELA 4: Modelo de Tabela de Contingência com Estrutura de Grupo nas duas Variáveis .... | 49 |
| TABELA 5: Tabela de Dados .....   | 70 |

## LISTA DE QUADROS

|   |    |
|---|----|
| QUADRO 1: Diagrama de Fluxo dos dados .....   | 76 |
| QUADRO 2: Inércias Principais dos Eixos ..... | 77 |
| QUADRO 3: Auxílio à descrição do eixo 1 ..... | 79 |
| QUADRO 4: Auxílio à descrição do eixo 2 ..... | 81 |

## LISTA DE GRÁFICOS

|   |    |
|---|----|
| GRÁFICO 1: Mapa da AC dos centróides dos grupos e variáveis descritoras ..... | 80 |
| GRÁFICO 2: Efeito feradura .....  | 83 |
| GRÁFICO 3: Mapa da AC com os pontos suplementares incluído .....              | 84 |
| GRÁFICO 4: Localização 5 .....  | 85 |
| GRÁFICO 5: Localização 4 .....  | 87 |
| GRÁFICO 6: Localização 3 .....  | 87 |
| GRÁFICO 7: Localização 2 .....  | 88 |
| GRÁFICO 8: Localização 1 .....  | 88 |

## LISTA DE NOTAÇÕES

|                      |   |
|----------------------|---|
| AC                   | ➤ Análise de Correspondência  |
| ACC                  | ➤ Análise de Correspondência Conjunta   |
| ACG                  | ➤ Análise de Correspondência Generalizada   |
| ACGr                 | ➤ Análise de Correspondência em Grupos  |
| ACM                  | ➤ Análise de Correspondência Múltipla   |
| ACP                  | ➤ Análise de Componentes Principais   |
| ACS                  | ➤ Análise de Correspondência Simples  |
| <b>C</b>             | ➤ nuvem (matriz) dos perfis colunas   |
| $C_j$                | ➤ $j$ -ésimo perfil coluna (vetor) de <b>C</b>  |
| $C_s$                | ➤ $s$ -ésimo perfil coluna (vetor) suplementar  |
| <b>c</b>             | ➤ vetor marginal das colunas de <b>P</b> ou vetor de massas associado à nuvem <b>C</b> ou centróide da nuvem <b>R</b> |
| $c_j$                | ➤ marginal da $j$ -ésima coluna de <b>P</b> ou massa associada ao $j$ -ésimo perfil coluna                            |
| $\bar{C}$            | ➤ centróide dos perfis coluna   |
| $\bar{c}$            | ➤ massa associada ao centróide dos perfis coluna  |
| $ca[R_i]_k$          | ➤ contribuição absoluta do $i$ -ésimo perfil linha para o $k$ -ésimo eixo   |
| $ca[C_j]_k$          | ➤ contribuição absoluta do $j$ -ésimo perfil coluna para o $k$ -ésimo eixo  |
| $cr[R_i]_k$          | ➤ contribuição relativa do $i$ -ésimo perfil linha para o $k$ -ésimo eixo   |
| $cr[C_j]_k$          | ➤ contribuição relativa do $j$ -ésimo perfil coluna para o $k$ -ésimo eixo  |
| $(C_j^\ell)$         | ➤ $j$ -ésimo perfil coluna do $\ell$ -ésimo grupo da matriz <b>N</b>  |
| $(c_j^\ell)$         | ➤ massa associada ao vetor $(C_j^\ell)$   |
| $(C_\bullet^\ell)_1$ | ➤ $\ell$ -ésimo perfil coluna da matriz <b>N</b> <sub>1</sub>   |
| $(c_\bullet^\ell)_1$ | ➤ massa associada ao vetor $(C_\bullet^\ell)_1$   |
| $(C_j^\ell)_2$       | ➤ $j$ -ésimo perfil coluna do $\ell$ -ésimo grupo da matriz <b>N</b> <sub>2</sub>                                     |

- $(c_j^l)_2$  ➤ massa associada ao vetor  $(C_j^l)_2$
- $(C_\bullet^l)_3$  ➤  $l$ -ésimo perfil coluna de  $N_3$
- $(c_\bullet^l)_3$  ➤ massa associada ao vetor  $(C_\bullet^l)_3$
- $D_c$  ➤ matriz diagonal cuja diagonal é o vetor  $c$
- $D_c^{-1}$  ➤ matriz inversa de  $D_c$
- $D_c^{-1/2}$  ➤ matriz diagonal cuja diagonal é formada pelo inverso da raiz quadrada dos elementos de  $c$
- $D_{\bar{c}}$  ➤ matriz diagonal cuja diagonal é o vetor  $\bar{c}$
- $D_{\bar{c}}^{-1}$  ➤ matriz inversa de  $D_{\bar{c}}$
- $D_{\bar{c}}^{-1/2}$  ➤ matriz diagonal cuja diagonal é formada pelo inverso da raiz quadrada dos elementos de  $\bar{c}$
- $D_r$  ➤ matriz diagonal cuja diagonal é o vetor  $r$
- $D_r^{-1}$  ➤ matriz inversa da  $D_r$
- $D_r^{-1/2}$  ➤ matriz diagonal cuja diagonal é formada pelo inverso da raiz quadrada dos elementos de  $r$
- $D_{\bar{r}}$  ➤ matriz diagonal cuja diagonal é o vetor  $\bar{r}$
- $D_{\bar{r}}^{-1}$  ➤ matriz inversa de  $D_{\bar{r}}$
- $D_{\bar{r}}^{-1/2}$  ➤ matriz diagonal cuja diagonal é formada pelo inverso da raiz quadrada dos elementos de  $\bar{r}$
- $D_\lambda$  ➤ matriz diagonal dos autovalores associado aos autovetores de  $S$
- $D_\lambda^{1/2}$  ➤ matriz diagonal dos valores singulares de  $X$
- $D_\eta$  ➤ matriz diagonal dos autovalores associado aos autovetores de  $Z$
- $d^2(R_i, R_i)$  ➤ distância qui-quadrado entre dois perfis definida pela matriz diagonal onde na diagonal estão os elementos do centróide da nuvem  $R$
- $F$  ➤ matriz das coordenadas principais dos perfis linha
- $F_i$  ➤ vetor das coordenadas principais do  $i$ -ésimo perfil linha
- $F_s$  ➤ vetor das coordenadas principais do  $s$ -ésimo perfil linha suplementar
- $f_{ik}$  ➤ coordenada principal (projeção) do  $i$ -ésimo perfil linha no  $k$ -ésimo eixo
- $G$  ➤ matriz das coordenadas principais dos perfis coluna

|                            |  |
|----------------------------|--|
| $G_s$                      | ➤ vetor das coordenadas principais do $s$ -ésimo perfil coluna suplementar   |
| $I$                        | ➤ total de linhas da TC ou $n^\circ$ de categorias da variável A   |
| $I-1$                      | ➤ dimensão do espaço dos perfis linha  |
| $I_h$                      | ➤ número de categorias no $h$ -ésimo grupo das linhas  |
| $In_{\bar{R}}(R_i)$        | ➤ inércia do $i$ -ésimo perfil linha com relação ao centróide de $\mathbf{R}$  |
| $In_{\bar{R}}(\mathbf{R})$ | ➤ inércia total dos perfis linha em relação ao centróide de $\mathbf{R}$   |
| $In_y(R_i)$                | ➤ inércia do $i$ -ésimo perfil linha em relação a um ponto qualquer $\mathbf{y} \in \mathbf{R}^J$  |
| $In_y(\mathbf{R})$         | ➤ inércia total dos perfis linha em relação a um ponto qualquer $\mathbf{y} \in \mathbf{R}^J$  |
| $In_y(\bar{\mathbf{R}})$   | ➤ inércia do centróide $\bar{\mathbf{R}}$ em relação a um ponto qualquer $\mathbf{y} \in \mathbf{R}^J$   |
| $In_{\bar{C}}(C_j)$        | ➤ inércia do $j$ -ésimo perfil coluna com relação ao centróide de $\mathbf{C}$   |
| $In_{\bar{C}}(\mathbf{C})$ | ➤ inércia total dos perfis coluna em relação ao centróide de $\mathbf{C}$  |
| $In_y(C_j)$                | ➤ inércia do $j$ -ésimo perfil coluna em relação a um ponto qualquer $\mathbf{y} \in \mathbf{R}^I$   |
| $In_y(\mathbf{C})$         | ➤ inércia total dos perfis colunas em relação a um ponto qualquer $\mathbf{y} \in \mathbf{R}^I$ .  |
| $[In_c(R_i)]_k$            | ➤ componente da inércia do $i$ -ésimo perfil linha com relação ao centróide de $\mathbf{R}$ no $k$ -ésimo eixo   |
| $J$                        | ➤ total de colunas da TC ou $n^\circ$ de categorias da variável B  |
| $J-1$                      | ➤ dimensão do espaço dos perfis coluna   |
| $J_\ell$                   | ➤ número de categorias no $\ell$ -ésimo grupo das colunas.   |
| $K$                        | ➤ dimensão do subespaço ajustado   |
| $K^*$                      | ➤ dimensão do subespaço ótimo  |
| $L(\mathbf{x}, \lambda_1)$ | ➤ equação de Lagrange onde $\lambda_1$ é o multiplicador de Lagrange   |
| $\mathbf{N}$               | ➤ matriz de dados associado à tabela de contingência ou matriz das frequências absolutas   |
| $\mathbf{N}_1$             | ➤ matriz das frequências das $I$ linhas nos $L$ grupos das colunas   |
| $\mathbf{N}_2$             | ➤ matriz das frequências dos $H$ grupos das linhas nas $J$ colunas   |
| $\mathbf{N}_3$             | ➤ matriz das frequências dos $H$ grupos das linhas nos $L$ grupos das colunas  |
| $N_i$                      | ➤ ponto cujas coordenadas são os elementos da $i$ -ésima linha de $\mathbf{N}$   |
| $n$                        | ➤ número total de casos observados na TC   |
| $n_{ij}$                   | ➤ número de casos observados da $i$ -ésima categoria da variável A e da $j$ -ésima categoria da variável B da TC ou interseção da $i$ -ésima linha e $j$ -ésima coluna de $\mathbf{N}$ |

- $n_{i\bullet}$  ➤ marginais das linhas da TC
- $n_{\bullet j}$  ➤ marginais das colunas da TC
- $n_{\bullet\bullet} = n$  ➤ número total de casos observados na TC
- $n_{ij}^{h\ell}$  ➤ número de indivíduos na  $i$ -ésima categoria do  $h$ -ésimo grupo da variável categórica A e na  $j$ -categoria do  $\ell$ -ésimo grupo da variável categórica B da TC ou interseção da  $i$ -ésima linha do  $h$ -ésimo grupo das linhas com a  $j$ -ésima coluna do  $\ell$ -ésimo grupo das colunas de **N**
- $n_{\bullet j}^{h\ell}$  ➤ número total de indivíduos do  $h$ -ésimo grupo da variável categórica A, na  $j$ -ésima categoria do  $\ell$ -ésimo grupo da variável categórica B.
- $n_i^{h\ell}$  ➤ número total de indivíduos do  $\ell$ -ésimo grupo da variável categórica B, na  $i$ -ésima categoria do  $h$ -ésimo grupo da variável categórica A.
- $n_{\bullet\bullet}^{j\ell}$  ➤ número total de indivíduos da variável categórica A, na  $j$ -ésima categoria do  $\ell$ -ésimo grupo da variável categórica B
- $n_i^{h\bullet}$  ➤ número total de indivíduos da variável categórica B, na  $i$ -ésima categoria do  $h$ -ésimo grupo da variável categórica A
- $n_{\bullet\bullet}^{h\ell}$  ➤ número total de indivíduos do  $h$ -ésimo grupo da variável categórica A, no  $\ell$ -ésimo grupo da variável categórica B
- $n_{\bullet\bullet}^{h\bullet}$  ➤ número total de indivíduos do  $h$ -ésimo grupo da variável categórica A, na variável categórica B
- $n_{\bullet\bullet}^{\bullet\ell}$  ➤ número total de indivíduos do  $\ell$ -ésimo grupo da variável categórica B na variável categórica A
- $n_{\bullet\bullet\bullet} = n$  ➤ número total de indivíduos da tabela
- P** ➤ matriz de correspondência
- P'** ➤ matriz transposta de **P**
- $p_{ij}$  ➤ proporção de casos observado na  $i$ -ésima linha e  $j$ -ésima coluna da TC ou interseção da  $i$ -ésima linha com a  $j$ -ésima coluna de **P**
- $qr[\mathbf{R}_i]$  ➤ qualidade de representação do  $i$ -ésimo perfil linha

|                     |  |
|---------------------|--|
| $qr[C_j]$           | ➤ qualidade de representação do $j$ -ésimo perfil coluna   |
| $\mathbf{R}^J$      | ➤ espaço vetorial $J$ -dimensional estruturado pela métrica qui-quadrado definida pela matriz diagonal onde na diagonal estão o inverso dos elementos do centróide dos perfis linha  |
| $\mathbf{R}^I$      | ➤ espaço vetorial $I$ -dimensional estruturado pela métrica qui-quadrado definida pela matriz diagonal onde na diagonal estão o inverso dos elementos do centróide dos perfis coluna |
| $\mathbf{R}$        | ➤ matriz dos perfis linhas (nuvem de pontos perfis linha)  |
| $R_i$               | ➤ $i$ -ésimo perfil linha de $\mathbf{R}$ (perfil da $i$ -ésima linha de $\mathbf{N}$ )  |
| $R_{ij}$            | ➤ interseção da $i$ -ésima linha e $j$ -ésima coluna de $\mathbf{R}$   |
| $R_s$               | ➤ perfil da $s$ -ésima linha suplementar   |
| $\mathbf{r}$        | ➤ vetor marginal das linhas de $\mathbf{P}$ ou vetor de massas das linhas de $\mathbf{R}$ ou centróide da nuvem $\mathbf{C}$   |
| $r_i$               | ➤ marginais das linhas da matriz de correspondência $\mathbf{P}$ ou massa do $i$ -ésimo perfil linha   |
| $\bar{\mathbf{R}}$  | ➤ centróide dos perfis linha ( $\mathbf{R}$ )  |
| $\bar{r}$           | ➤ massa de $\bar{\mathbf{R}}$  |
| $(R_i^h)_1$         | ➤ perfil linha da $i$ -ésima linha do $h$ -ésimo grupo de $\mathbf{N}_1$   |
| $(r_i^h)_1$         | ➤ massas dos perfis linhas de $\mathbf{N}_1$   |
| $(R_{\bullet}^h)_2$ | ➤ perfil da $h$ -ésima linha de $\mathbf{N}_2$   |
| $(r_{\bullet}^h)_2$ | ➤ massas dos perfis da $h$ -ésima linha de $\mathbf{N}_2$  |
| $(R_{\bullet}^h)_3$ | ➤ perfis linhas de $\mathbf{N}_3$  |
| $(r_{\bullet}^h)_3$ | ➤ massas dos perfis linhas de $\mathbf{N}_3$   |
| $\mathbf{S}$        | ➤ matriz que define o ajuste do subespaço dos perfis linha   |
| $\mathbf{S}^*$      | ➤ matriz de inércias dos perfis linha não centralizados  |
| TC                  | ➤ Tabela de Contingência   |
| $tr[.]$             | ➤ traço de uma matriz  |
| $\mathbf{U}$        | ➤ matriz dos autovetores de $\mathbf{Z}$   |
| $\mathbf{V}$        | ➤ matriz dos autovetores de $\mathbf{S}$   |

- $v$  ➤ vetores direcionais unitário cuja direção mostre a maior inércia dos pontos da nuvem
- $v_k$  ➤ vetores direcionais do  $k$ -ésimo eixos principais
- $v_x$  ➤ vetor tal que  $v_x v_x = 1$
- $v_a$  ➤ autovetor de  $S$  associado ao autovalor 0 e autovetor de  $S^*$  associado ao autovalor 1
- $X$  ➤ matriz dos perfis linha ponderados, transformados e centralizados
- $X^*$  ➤ matriz dos perfis linha ponderados e transformados
- $Y$  ➤ matriz dos perfis coluna ponderados, transformados e centralizados
- $Y^*$  ➤ matriz dos perfis coluna ponderados e transformados
- $Z$  ➤ matriz que define o ajuste do subespaço dos perfis coluna
- $\lambda_k$  ➤ inércia do  $k$ -ésimo eixos principais (indica o quanto a
- $\chi^2$  ➤ estatística qui-quadrado que testa a independência entre linhas e colunas de uma TC
- $\Phi$  ➤ matriz das coordenadas padronizadas dos perfis linha
- $\Gamma$  ➤ matriz das coordenadas padronizadas dos perfis coluna
- $\Gamma_j$  ➤ vetor de coordenadas padronizadas do  $j$ -ésimo perfil coluna
- $\theta_k$  ➤ ângulo entre o vetor direcional do ponto  $R_i$  e o eixo  $v_k$
- $\nabla$  ➤ gradiente (vetor de derivadas)
- $\mathbf{1}_i$  ➤ vetor com  $i$  componentes todos iguais a um
- $\mathbf{1}_j$  ➤ vetor com  $j$  componentes todos iguais a um

## RESUMO

Neste trabalho, a técnica de análise multivariada para dados categóricos, Análise de Correspondência (AC) é discutida de forma essencialmente geométrica e descritiva. É feita uma listagem dos vários tipos de dados que podem ser utilizados como entrada desse algoritmo, sendo a tabela de contingência a forma mais simples. Deste modo, desenvolvemos a Análise de Correspondência Simples (ACS), que é o estudo da Análise de Correspondência para tabelas de contingência simples, com a ênfase usual da visão geométrica.

A Análise de Correspondência para tabelas de contingência que apresentam uma estrutura de grupo, que chamamos de *Análise de Correspondência em Grupos* (ACGr) é, então, discutida como uma extensão da ACS, onde o interesse é o de conhecer as associações entre os agrupamentos das categorias nesta estrutura, mas também conhecer as associações das categorias dentro de cada grupo e entre os grupos. Assim, podemos dizer que a ACGr compartilha o princípio da Análise Discriminante Canônica de Fisher: o de maximizar a variabilidade entre os grupos, aqui inércia entre grupos. A ACGr pode ser estruturada em três etapas. A primeira é definir um representante para cada grupo, aqui definimos o centróide como o representante. A segunda é maximizar a inércia destes representantes, que é equivalente a desenvolver a ACS para as tabelas de contingências formadas pelos representantes dos grupos. A terceira é colocar todas as categorias dos grupos como pontos suplementares na estrutura definida pelos centróides. Para validar esta estrutura mostramos as relações entre as várias matrizes associadas a uma tabela de contingência com estrutura de grupo e demonstramos os principais resultados, como enunciados em Greenacre(1984).

Uma aplicação da ACGr é feita trabalhando com dados dos censos de 1970, 1980 e 1991, onde o interesse é analisar a associação entre variáveis descritoras (Instrução, Renda e Localização do Domicílio) e a variável resposta (Localização Doméstica) controlada pela variável Ciclos de Vida.

# 1. INTRODUÇÃO

## 1.1 APRESENTAÇÃO

Muitas vezes torna-se necessário conhecermos mais profundamente alguns fenômenos e, sem dúvida, a estatística é uma poderosa aliada na busca deste conhecimento; seu potencial em resumir, simplificar e eventualmente explicar um conjunto de dados não pode ser desprezado. Hoje, tanto na física das partículas, quanto na arqueologia, na biologia, na economia, na medicina, na psicologia, entre outras áreas, existe, cada vez mais, a necessidade de se manipular uma imensa quantidade de dados, e isso, sem a estatística e sem os computadores que permitem a realização rápida de complicados cálculos matemáticos, seria muito mais difícil.

Utilizando-se tanto das teorias matemáticas quanto dos computadores, a estatística tem mostrado ser uma ferramenta poderosa para o entendimento dos fenômenos nas diversas áreas. Através da análise multivariada, ela tem apelado aos espaços matemáticos, que comportam um número qualquer de dimensões, para representar, resumir e visualizar o comportamento desses fenômenos. Em geral, as técnicas de análise de dados multivariados têm como objetivo fornecer uma representação sucinta de grandes massas de dados, que na maioria das vezes são de difícil interpretação.

Existe uma grande variedade de técnicas de análise de dados multivariados que visam reduzir a dimensionalidade dos dados através do cálculo dos autovalores e autovetores da matriz que os representa. Essa redução é para facilitar a interpretação dos dados na sua representação

gráfica. Dentre essas técnicas está a *Análise de Correspondência* (AC), que é aplicada principalmente aos dados cujas variáveis sejam categóricas nominais ou ordinais.

O uso desta técnica vem crescendo, com o respaldo da informática, nas diversas áreas da ciência, principalmente nos complexos problemas das ciências sociais. O poder de síntese dos resultados em termos estatísticos em muito acrescenta ao poder de síntese das áreas substantivas em que se presta como estratégia (Greenacre e Blasius, 1994).

Na manipulação de dados categóricos pode-se obter uma tabela de duas entradas que nos dá as freqüências absolutas dos cruzamentos das categorias de duas variáveis categóricas. Esta tabela é chamada *Tabela de Contingência* (TC). Quando usamos a TC como 'entrada' do algoritmo da AC, temos a forma mais simples do seu desenvolvimento e ainda obtemos as regras básicas para a sua interpretação geométrica.

Em algumas tabelas de contingência podem existir partições em suas categorias ou nos indivíduos que as compõem. Nestes casos as categorias podem ser agrupadas em classes. Para cada partição, estas classes, como as categorias de uma tabela de contingência, devem ser disjuntas, para que nenhuma categoria ou indivíduo pertença a mais de uma classe, e exaustivas, para que todas as categorias ou indivíduos sejam classificados. Sobre este tipo de tabelas dizemos que elas possuem uma *estrutura de grupo* ou tem uma sobre-estruturação nas categorias de suas variáveis. As tabelas com estrutura de grupo podem ter suas categorias agrupadas em classes nas linhas e/ou colunas, de acordo com os interesses do estudo nas associações existentes entre os agrupamentos dessas categorias. Conforme esses interesses, podemos obter pelo menos três novas tabelas a partir de uma tabela com estrutura de grupo. É para dados cujas tabelas de contingência tenham estrutura de grupo que este trabalho se apresenta.

Na próxima seção, a seção 1.2, damos uma visão geral do que é a AC, mostrando o quanto é ampla a aplicabilidade do seu algoritmo, a nomenclatura que ela recebe quando os vários tipos de dados são tomados como entrada desse algoritmo e um breve histórico dessa técnica.

No capítulo 2, mostramos o desenvolvimento do algoritmo da análise de correspondência para tabelas de contingência onde não se considera a estrutura de grupo. Tal desenvolvimento foi feito como uma generalização dupla, nas linhas e nas colunas, da técnica Análise de Componentes Principais. Esse desenvolvimento servirá de base teórica para o estudo das associações entre as categorias que tenham uma estrutura de grupo em suas variáveis.

No capítulo 3, fazemos este estudo para tabelas com apenas uma partição em cada variável. Neste caso, ao fazermos o agrupamento das linhas e/ou das colunas, obtemos três novas tabelas. Neste estudo, mostramos os resultados que relacionam as análises das quatro diferentes formas de visualização desta tabela de contingência a partir da consideração da estrutura de grupo. E ainda mostramos os resultados que relacionam os grupos dentro destas matrizes.

No capítulo 4, apresentamos uma aplicação da análise de correspondência para dados que tenham estrutura de grupo. Tal aplicação diz respeito à organização doméstica de pessoas sem companheiro no estado de São Paulo. O conjunto de dados foi obtido através dos Censos Demográficos de 1970, 1980 e 1991, e nos traz uma grande quantidade de informações que dão origem a grandes tabelas de contingência que podem ser agrupadas em classes. Deste modo, este conjunto de dados mostra-se adequado para a aplicação do nosso estudo.

Finalmente, no capítulo 5, fazemos algumas considerações sobre a técnica e sobre este trabalho. Apresentamos ainda algumas sugestões, para futuros trabalhos, que podem ser desenvolvidos a partir deste.

A relevância do presente trabalho está em colocar ao alcance dos pesquisadores aspectos mais poderosos da técnica em problemas com sobre-estrutura de linha e/ou coluna. Aos estatísticos, as derivações algébricas permitem o domínio do conteúdo da técnica e, aos usuários de outras áreas, o exemplo desenvolve-se de forma a ilustrar o valor agregado em síntese nos problemas deste tipo.

## **1.2 ANÁLISE DE CORRESPONDÊNCIA**

Nesta seção, damos uma visão geral do que é a AC. Falamos sobre a amplitude de aplicação do seu algoritmo em dados categorizados, sobre a nomenclatura que ela recebe quando os vários tipos de dados são tomados como entrada desse algoritmo e um pouco de sua história. É com o intuito de introduzir esta técnica que esta seção se apresenta.

### **1.2.1 A TÉCNICA**

A Análise de Correspondência (AC) é uma técnica multivariada para análise exploratória de dados categorizados. Ela converte uma matriz de dados não negativos em um particular tipo de gráfico que exhibe as linhas e colunas da matriz como pontos de um espaço vetorial de dimensão menor que a original, de maneira que as relações entre as linhas, entre as colunas e entre linhas e colunas possam ser interpretadas, como colocou Greenacre e Hastie (1987). É exatamente por estar interessada em estudar a correspondência entre variáveis, que esta técnica recebeu o nome de Análise de Correspondência.

Sua geometria e álgebra fazem com que pertença a uma família de técnicas de disposição gráfica que são baseadas em aproximação de uma matriz por outra de posto menor, por meio da decomposição em valores singulares. Algumas outras técnicas que fazem parte desta família são, por exemplo, a Análise de Componentes Principais, o “Biplot”, a Correlação Canônica, entre outras. É conveniente descrever a AC como uma generalização da análise de componentes principais, técnica introduzida por Pearson (1901). Assim podemos dizer que o problema da AC é achar um subespaço que “melhor ajuste” a nuvem de pontos no espaço euclidiano. Este ajuste é feito pelo método de quadrado mínimo ponderado onde é utilizada a distância euclidiana generalizada (ponderada) em um sistema de massas pontuais.

No contexto deste trabalho, ao discutir a análise de correspondência, nós o faremos tomando por base três etapas que se originaram das idéias de Pearson acerca da Análise de Componentes Principais, colocadas por Greenacre e Vrba (1984). Estas etapas são: a) definir duas nuvens de pontos em correspondência num espaço multidimensional, b) definir uma estrutura

métrica em cada nuvem de pontos, c) definir o ajuste de cada nuvem de pontos para um subespaço de dimensão menor que o original, tal que a representação dos pontos seja a mais precisa possível.

A abordagem da AC enfatiza a representação geométrica, revelando a estrutura dos dados de forma ótima e sem necessidade de suposição de modelo ou distribuição para os dados. Essa representação é de tal maneira que as distâncias entre os pontos linha (ou coluna) no espaço euclidiano são iguais à distância qui-quadrado - lembrando que essa representação geométrica só é visualizada em espaços vetoriais de dimensões menores ou iguais a três.

Podemos então dizer que um dos propósitos da AC é o de reduzir a dimensão do espaço vetorial, conservando da melhor forma possível a configuração inicial. Isto significa que podemos considerar a AC como sendo um algoritmo de redução de dados que fornece imagens simplificadas da realidade multidimensional e busca a melhor representação simultânea dos conjuntos de categorias da matriz de dados.

### 1.2.2 APLICAÇÃO

O algoritmo da AC tem uma aplicação muito ampla dentre os dados categorizados já que, como podemos observar na definição dada por Greenacre e Hastie (1987) transcrita no primeiro parágrafo da seção anterior, a única restrição é que os dados sejam não negativos. Como exemplo de dados onde podemos aplicar esse algoritmo estão aqueles contidos nas *Tabelas de Códigos Condensados*, denominação encontrada em Prado (1996) referente às tabelas provenientes de resumos codificados de respostas a uma enquete. Estas tabelas são compostas de linhas por indivíduos e colunas por característica observada, onde cada linha contém todos os códigos correspondentes às modalidades atribuídas a um indivíduo para cada uma das características observadas. Na interseção da  $i$ -ésima linha com a  $j$ -ésima coluna aparece o valor  $k_{ij}$  que corresponde ao código numérico que foi dado à modalidade atribuída ao  $i$ -ésimo indivíduo com relação a  $j$ -ésima característica observada. Este tipo de tabela é, também, muito comum em estudos de preferência, onde também se aplica este algoritmo, como podemos observar em Moisés Ortega (1997).

Quando se faz a justaposição de uma tabela de códigos condensados, onde estes códigos têm uma ordenação, com uma outra em cujas caselas correspondentes estão os códigos numéricos

complementares, obtemos a chamada *Tabela Dobrada* (Doubled Table, Greenacre(1993)) que também é usada como entrada do algoritmo.

Um outro exemplo é a chamada Tabela Lógica ou *Tabela Disjuntiva Completa*. Ela é correspondente à tabela de códigos condensados onde o valor do código numérico assume apenas 1 (um) ou 0 (zero) para indicar se o indivíduo tem ou não aquela característica.

O algoritmo é aplicado também às tabelas de duas entradas, às tabelas de mais de duas entradas e às tabelas de Burt. As tabelas de duas entradas são as chamadas *Tabelas de Contingência*. As tabelas de três ou mais entradas nos dão as frequências absolutas dos cruzamentos das categorias de três ou mais variáveis (conjunto de características ou de categorias), da tabela disjuntiva completa. Estas são chamadas de *Tabelas de Multi-entrada*. A *Tabela de Burt* é o cruzamento entre várias variáveis categóricas duas a duas. De outro modo, a tabela de Burt é constituída de um mosaico de tabelas onde as tabelas sobre a diagonal principal são matrizes diagonais e as demais são tabelas de contingência que são representadas por matrizes retangulares. A tabela de Burt é análoga à matriz de variâncias e covariâncias para o caso contínuo. Esse algoritmo ainda pode ser aplicado às variáveis contínuas desde que seja feita a devida categorização. Mas o que mais nos interessa neste trabalho é que esse algoritmo também pode ser aplicado à tabelas de contingência com estrutura de grupo.

O estudo da AC, ou seja, a aplicação desse algoritmo às diferentes tabelas de entrada, recebe denominações distintas. Por exemplo, o estudo da AC em tabelas de contingência simples é denominado *Análise de Correspondência Simples* (ACS). Olhando ACS como uma técnica que estuda as TC, podemos dizer que ela trabalha com a diferença entre as observações de uma tabela de contingência e a esperança sob a hipótese de independência entre as variáveis categóricas. Sob esta ótica, o estudo da AC em tabelas de contingência onde se trabalha com a diferença entre as observações e a esperança sob algum modelo geral é denominado *Análise de Correspondência Generalizada* (ACG). Por outro lado, o estudo da AC em tabelas de contingência onde existe uma estrutura de grupo nas variáveis, é denominado *Análise de Correspondência em Grupos* (ACGr), aqui a ótica é a da análise discriminante. Ainda o estudo da AC em tabelas de contingência justapostas, onde uma das variáveis é igual, é denominado *Análise de Correspondência Conjunta* (ACC). O estudo da AC em tabela de Burt ou em tabelas lógicas é denominado *Análise de Correspondência Múltipla* (ACM). Todas estas análises investigam as associações entre as

variáveis categóricas duas a duas. A investigação das associações entre três ou mais variáveis categóricas é discutida rapidamente por Greenacre (1993) quando se estuda a AC em tabelas de multi-entrada.

### 1.2.3 UM BREVE HISTÓRICO

A AC é conhecida, algebricamente, há mais de 60 anos; ela teve seus primeiros cálculos matemáticos feitos por Hirschfeld em 1935 e Fisher em 1940, como é ressaltado em Hill (1974). Por algum tempo o mesmo procedimento algébrico foi utilizado em diferentes contextos, recebendo diferentes denominações - Ponderação Recíproca, 'Dual ou Optimal Scaling', Correlação Canônica, Regressão Linear Simultânea, Análise de Homogeneidade, entre outras - dependendo da área de onde os dados eram originários.

Esta técnica, por muito tempo negligenciada, foi retomada na França por um grupo de pesquisadores liderado por Jean-Paul Benzécri, em meados da década de 60, como uma técnica gráfica de análise de dados. Inicialmente teve a denominação Análise Fatorial de Correspondência e depois simplesmente Análise de Correspondência. A partir daí uma extensa quantidade de literatura sobre o assunto vem sendo gerada.

Hoje a técnica está difundida pelo mundo todo, sendo a África do Sul, Europa e Japão seus principais centros de pesquisa. Na Europa, principalmente na França, destacam-se vários colaboradores do grupo de J-P Benzécri. Na África do Sul destaca-se o M. J. Greenacre, que foi orientado no seu doutorado na França por J-P Benzécri. E no Japão destaca-se S. Nishisato que vem de uma outra escola. Esta técnica tem hoje aplicações em diversas áreas de pesquisa tais como: sociologia, ecologia, paleontologia, arqueologia, geologia, educação, medicina, lingüística, pesquisa de mercado, ciências agrárias, estudos religiosos, arte, música, epidemiologia e outras, mediante grupos interdisciplinares em universidades e/ou unidades de pesquisa.

## 2. ANÁLISE DE CORRESPONDÊNCIA SIMPLES

### 2.1 INTRODUÇÃO

A Análise de Correspondência Simples (ACS) é o nome que se dá ao estudo da AC quando os dados são apresentados sob a forma de tabela de contingência. Esse estudo nos mostra, de forma mais simples, o desenvolvimento do algoritmo da AC e, além disso, sua geometria fornece as regras básicas para a interpretação. Todas as outras formas de análise de correspondência são aplicações do mesmo algoritmo em outros tipos de matriz de dados, com uma conseqüente adaptação da interpretação. A ACS nos fornece a melhor representação simultânea das duas variáveis categóricas que são constituídas pelas linhas e colunas da matriz de dados, apresentando gráficos onde aparecem as nuvens dos pontos linha e dos pontos coluna no espaço projeção de dimensão desejável menor ou igual a três. Isso é o que apresentamos nesse capítulo.

Para desenvolvermos a ACS, na seção 2.2, mostramos como os dados são apresentados e as matrizes que os representam.

Na seção 2.3, a partir da representação das linhas da matriz de dados, definimos uma nuvem de pontos no espaço vetorial, a massa de cada ponto dessa nuvem e o centro de gravidade da nuvem com sua respectiva massa. Definimos a métrica qui-quadrado como a adequada para a AC. A partir daí, definimos a inércia dos pontos e a transformação desses pontos para melhor compreensão do gráfico. Definimos ainda o ajuste da nuvem a um subespaço ótimo, dando as coordenadas dos pontos nesse espaço.

Na seção 2.4 mostramos as relações entre os pontos, vindas da representação das colunas da matriz de dados, através de um estudo análogo ao da seção anterior, passando pelas mesmas etapas mas, agora, apresentando apenas os resultados.

Após termos mostrado as relações entre os pontos linha (seção 2.3) e as relações entre os pontos coluna (seção 2.4), mostramos na seção 2.5, as relações entre os pontos linha e os pontos coluna, bem como a correspondência entre as duas nuvens de pontos. Ao representarmos as duas nuvens num mesmo sistema de eixos, observamos as associações entre as linhas e as colunas da matriz de dados, através das fórmulas de transição.

Alguns autores -ver Villalobos Aguaio (1993)- consideram, arte, uma parte da interpretação da Análise de Correspondência. Neste trabalho, procuramos facilitar o caminho aos necessários “insights”, com o intuito de socializar o uso da técnica. Para tanto, recuperamos e aprimoramos um roteiro que poderá ser utilizado na grande maioria dos casos de aplicação da AC. Assim, na seção 2.6, que finaliza este capítulo, apresentamos algumas considerações para interpretar os gráficos cujos pontos são as projeções dos perfis sobre o subespaço ótimo. Também nesta seção apresentamos os pontos suplementares e sua caracterização. Estes pontos têm um papel muito importante na análise dos dados com estrutura de grupo.

## 2.2 OS DADOS

Uma etapa indispensável do tratamento de um conjunto de dados de um problema é a sua transcrição numérica, e aí temos várias possibilidades de codificação a partir dos quais podemos construir diferentes tipos de tabelas. Uma delas é a tabela de frequências absolutas chamada Tabela de Contingência. Os dados onde se aplicam a ACS são apresentados em uma tabela de contingência onde, nas linhas, figuram  $I$  categorias de uma variável categórica A e, nas colunas,  $J$  categorias de uma outra variável categórica B. Na interseção das linhas e colunas, o número  $n_{ij}$ ,  $i = 1, \dots, I$  e  $j = 1, \dots, J$ , é o número de casos observados da  $i$ -ésima categoria da primeira variável A e da  $j$ -ésima categoria da segunda variável B.

A cada tabela associamos uma matriz, através da qual representamos os dados e aplicamos as teorias matemáticas. Assim, no caso de uma tabela de contingência (Tabela 1), associamos a matriz  $N_{I \times J}$ , dada em (2.2.1), denominada **Matriz de Dados**. Assumindo-as como sinônimas, quando nos referirmos à matriz N estaremos falando da tabela 1 e vice versa.

TABELA 1: Modelo de Tabela de Contingência Simples.

|                             |             | Variável Categórica B |     |                 |                      |
|-----------------------------|-------------|-----------------------|-----|-----------------|----------------------|
|                             |             | Categoria 1           | ... | Categoria J     | Total                |
| Variável<br>Categórica<br>A | Categoria 1 | $n_{11}$              | ... | $n_{1J}$        | $n_{1\bullet}$       |
|                             | ⋮           | ⋮                     |     | ⋮               | ⋮                    |
|                             | Categoria I | $n_{I1}$              | ... | $n_{IJ}$        | $n_{I\bullet}$       |
|                             | Total       | $n_{\bullet 1}$       | ... | $n_{\bullet J}$ | $n_{\bullet\bullet}$ |

$$N = \begin{bmatrix} n_{11} & \cdots & n_{1J} \\ \vdots & & \vdots \\ n_{I1} & \cdots & n_{IJ} \end{bmatrix} = [n_{ij}]_{I \times J}, n_{ij} \geq 0. \quad (2.2.1)$$

Denotaremos as marginais das linhas e das colunas de N, respectivamente, por

$$n_{i\bullet} = \sum_{j=1}^J n_{ij} > 0, i = 1, \dots, I \quad (2.2.2)$$

$$n_{.j} = \sum_{i=1}^I n_{ij} > 0, j=1, \dots, J \quad (2.2.3)$$

e o total geral por

$$n_{..} = \sum_{i=1}^I \sum_{j=1}^J n_{ij} = n.$$

Obs. 1: Nos casos em que  $n_{i.}$  ou  $n_{.j}$  forem nulos, as categorias correspondentes devem ser retiradas da análise.

A partir da Matriz de Dados, matriz de freqüências absolutas, podemos construir uma outra matriz, uma matriz de freqüências relativas. O uso desta matriz no desenvolvimento da AC tem como objetivo simplificar a notação nas expressões matemáticas que nos dão a correspondência entre linhas e colunas da matriz de dados. Transformemos a matriz  $\mathbf{N}$ , mantendo as relações proporcionais entre as linhas e as colunas, dividindo cada elemento  $n_{ij}$  pelo total geral  $n$

e denotando seus elementos por  $p_{ij} = \frac{n_{ij}}{n}$ . Assim temos a matriz

$$\mathbf{P} = \frac{1}{n} \mathbf{N} = [p_{ij}]_{I \times J}, 0 \leq p_{ij} \leq 1 \quad (2.2.4)$$

denominada *Matriz de Correspondência*. As marginais das linhas e colunas da matriz  $\mathbf{P}$  são denotadas, respectivamente, por

$$r_i = \sum_{j=1}^J p_{ij}, i = 1, \dots, I,$$

$$c_j = \sum_{i=1}^I p_{ij}, j = 1, \dots, J.$$

Note que as marginais das linhas e das colunas da matriz  $\mathbf{P}$  são iguais às respectivas marginais das linhas e das colunas de  $\mathbf{N}$  divididas pelo total geral, isto é,

$$r_i = \frac{n_{i.}}{n}, i = 1, \dots, I,$$

$$c_j = \frac{n_{.j}}{n}, j = 1, \dots, J$$

e o total geral de  $\mathbf{P}$  é 1. Assim os vetores marginais de  $\mathbf{P}$  nos dão as freqüências relativas das linhas e das colunas. E os vetores marginais são dados, respectivamente, por:

$$\mathbf{r} = (r_1, \dots, r_I)^T,$$

$$\mathbf{c} = (c_1, \dots, c_J)^T.$$

Matricialmente esses vetores são expressos por

$$\mathbf{r} = \mathbf{P}\mathbf{1}_j, \quad (2.2.5)$$

$$\mathbf{c} = \mathbf{P}'\mathbf{1}_i \quad (2.2.6)$$

onde  $\mathbf{1}_j$  e  $\mathbf{1}_i$  são vetores com  $j$  e  $i$  componentes todos iguais a um (1).

Feita a apresentação dos dados, passemos a estudar as associações entre as categorias das variáveis em questão.

### 2.3 RELAÇÕES ENTRE AS LINHAS

Nesta seção mostramos como as linhas da matriz  $N$  se relacionam. Para isso, estas linhas foram representadas como pontos, num espaço vetorial, por seus perfis ao longo das colunas. Isto foi feito segundo as idéias de Pearson, isto é, definimos uma nuvem de pontos, a métrica para esta nuvem e o ajuste do subespaço ótimo. Isto é análogo a aplicar a técnica Análise de Componentes Principais a um conjunto de  $I$  observações e  $J$  variáveis, com uma generalização simples no uso de uma métrica generalizada e na ponderação das observações.

#### 2.3.1 A NUVEM DE PONTOS NO ESPAÇO VETORIAL $\mathbf{R}^J$

Ao se analisar um conjunto de dados e utilizar a Análise de Correspondência como instrumento, faz-se uso de gráficos para visualizar as relações entre as variáveis que se deseja analisar. Ao fazê-lo, parte-se da representação multidimensional, onde se encontra uma nuvem de pontos. Essa nuvem de  $I$  pontos é definida no espaço  $J$ -dimensional,  $\mathbf{R}^J$ , proveniente das  $I$  linhas da matriz de dados  $N$ . Ao representá-la não nos interessa ver as freqüências absolutas de cada linha, mas sim o perfil de uma linha ao longo das categorias das colunas. Assim representamos a  $i$ -ésima linha de  $N$  no espaço  $\mathbf{R}^J$  por seu perfil ao longo das colunas, que é o ponto do espaço  $\mathbf{R}^J$  dado por

$$R_i = \left( \frac{n_{i1}}{n_{i\bullet}} \quad \dots \quad \frac{n_{iJ}}{n_{i\bullet}} \right), \quad i=1, \dots, I \quad (2.3.1)$$

denominado o  $i$ -ésimo **perfil linha** de  $N$ , onde  $n_{i\bullet} \neq 0$  conforme restrição de (2.2.2) e Obs.1:.

Mas ao introduzir os perfis, padronizamos os pontos, perdendo assim a informação da proporção de indivíduos em cada linha. Daí a necessidade de atribuir uma massa a cada ponto. Essa massa é a freqüência relativa de cada linha da matriz  $N$ . Portanto, a **massa** do  $i$ -ésimo perfil linha é dada por

$$r_i = \frac{n_{i\bullet}}{n}, \quad i = 1, \dots, I. \quad (2.3.2)$$

Essas massas estão intimamente ligadas aos perfis, não podendo analisar os perfis sem deixar de considerá-las. É comum encontrar estudos, principalmente em tabelas pequenas, em que o

pesquisador transforma os dados em perfis (percentis com relação às marginais) e as vezes tira conclusões errôneas, dando-as como corretas. Isto acontece quando as proporções de indivíduos nas linhas ou colunas são diferentes.

Na figura 1 abaixo, mostramos a representação da  $i$ -ésima linha de  $\mathbf{N}$  e de seu perfil no espaço  $\mathbf{R}^3$ .

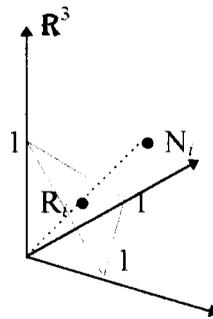


Figura 1: Representação do Perfil Linha  $R_i$  no e Espaço  $\mathbf{R}^3$

O fato de existir uma dependência linear entre as coordenadas dos perfis -isto é, de a soma das coordenadas ser igual a 1- significa geometricamente que esses pontos estão contidos numa região do espaço  $\mathbf{R}^{J-1}$  delimitada pelos segmentos de retas que ligam os pontos unitários de cada um dos eixos coordenados. Esta região é denominada “simplex” regular ou espaço dos perfis linha. Como se pode observar na figura 1 acima, para  $J=3$ , o “simplex” é um triângulo fechado pelos segmentos de retas que unem os três vértices unitários sobre os eixos perpendiculares.

Voltando ao caso do  $\mathbf{R}^J$ , percebemos que cada linha da matriz  $\mathbf{N}$  pode ser representada por um ponto do “simplex”. Assim, ao representar todas estas linhas, é formada uma nuvem de  $I$  pontos no “simplex”. Essa nuvem de pontos será, então, expressa por uma matriz  $I \times J$  dada por

$$\mathbf{R} = \begin{bmatrix} \frac{n_{11}}{n_{1\bullet}} & \dots & \frac{n_{1J}}{n_{1\bullet}} \\ \vdots & & \vdots \\ \frac{n_{I1}}{n_{I\bullet}} & \dots & \frac{n_{IJ}}{n_{I\bullet}} \end{bmatrix} = [R_1' | \dots | R_I'] = \mathbf{D}_r^{-1} \mathbf{P} \quad (2.3.3)$$

denominada *matriz dos perfis linha*. Como cada ponto tem associada uma massa, a nuvem de pontos, representada pela matriz  $\mathbf{R}$ , terá associado um vetor de massas, que nos dá as frequências relativas das linhas de  $\mathbf{N}$ . Ele é dado por

$$\mathbf{r} = \left( \frac{n_{1\bullet}}{n} \quad \dots \quad \frac{n_{I\bullet}}{n} \right), \quad (2.3.4)$$

e é denominado *vetor de massas* dos perfis linha. Assim, o vetor marginal  $\mathbf{r}$  definido em (2.2.5) é o vetor cujas coordenadas são as massas dos pontos da nuvem  $\mathbf{R}$ .

Um aspecto importante, no estudo das associações entre as categorias da variável (A) das linhas, é a dispersão dos pontos da nuvem  $\mathbf{R}$ . Essa dispersão é medida em relação a um ponto central da nuvem. Esse ponto é denominado *centróide* dos perfis linha ou centro de gravidade da nuvem dos perfis linha. Ele nos mostra o perfil da nuvem em relação às colunas, isto é, ele é o perfil da linha formado pela soma de todas as linhas de  $\mathbf{N}$ . Denotado por  $\bar{\mathbf{R}}$ , ele é dado por

$$\bar{\mathbf{R}} = \left( \frac{n_{\bullet 1}}{n} \quad \dots \quad \frac{n_{\bullet J}}{n} \right), \quad (2.3.5)$$

e sua massa é igual à soma das massas de todos os pontos da nuvem, que será denotada por  $\bar{r}$ , e dada por

$$\bar{r} = \sum_{i=1}^I r_i. \quad (2.3.6)$$

O centróide é um ponto que pode não ser um componente da nuvem dos perfis linha, mas com certeza é o candidato natural a representante desta nuvem; a qualidade desta representação dependerá obviamente da homogeneidade dos perfis linha. O centróide é o perfil médio da nuvem, mas uma média em que cada ponto  $\mathbf{R}_i$  é ponderado por sua massa  $r_i$ . Isso é uma generalização da -bem conhecida- noção de média, isto é,

$$\bar{\mathbf{R}} = \frac{1}{\bar{r}} \sum_{i=1}^I r_i \mathbf{R}_i = \mathbf{R}' \mathbf{r}. \quad (2.3.7)$$

Para tabelas de contingência simples (como a tabela 1) temos que  $\bar{r}=1$ , assim o centróide é igual à soma dos pontos  $\mathbf{R}_i$ ,  $i=1, \dots, I$ , ponderada pelas suas respectivas massas  $r_i$ .

Desta forma, definimos uma das nuvem de pontos com as suas respectivas massas; na próxima subseção vamos definir a métrica para essa nuvem de pontos.

### 2.3.2 A MÉTRICA DA NUVEM NO ESPAÇO $\mathbf{R}^J$

O formato da nuvem ou o aparecimento de pontos isolados pode indicar a existência ou não de associação entre os pontos. Essas associações são medidas através das distâncias entre os pontos. Daí a necessidade de se definir uma estrutura métrica para medir essas distâncias. Estruturaremos, então, o espaço  $\mathbf{R}^J$ , isto é, a nuvem dos perfis linha, de forma a podermos calcular distâncias entre os pontos dessa nuvem e também medir a dispersão da nuvem. Em Análise de Correspondência o estudo das nuvens é feito tomando como base o seu centróide ou o seu representante, daí a introdução de uma distância peculiar a esta técnica. Ela é uma distância euclidiana ponderada, onde a matriz que define a ponderação,  $\mathbf{D}_{\bar{\mathbf{R}}}^{-1}$ , é uma matriz diagonal em que, na diagonal, aparece o inverso dos elementos (coordenadas) do centróide. Esta distância é expressa por:

$$d^2(\mathbf{R}_i, \bar{\mathbf{R}}) = \sum_{j=1}^J \left( \frac{n_{ij}}{n_{i\cdot}} - \frac{n_{\cdot j}}{n} \right)^2 \frac{n}{n_{\cdot j}} = (\mathbf{R}_i - \bar{\mathbf{R}})' \mathbf{D}_{\bar{\mathbf{R}}}^{-1} (\mathbf{R}_i - \bar{\mathbf{R}}) \quad (2.3.8)$$

e é chamada de *Distância Qui-quadrado* ( $\chi^2$ ). Assim nuvens que têm o mesmo centróide têm definida a mesma distância. Esta métrica procura dar uma estabilidade aos pontos, já que cada termo do somatório que está ao quadrado em (2.3.8) é a contribuição da coluna  $j$  ao valor da distância entre os pontos  $\mathbf{R}_i$  e  $\bar{\mathbf{R}}$ , e o termo de ponderação  $\frac{n}{n_{\cdot j}}$  é introduzido para eliminar a influência das colunas que possuem marginais com valores absolutos muito altos com relação às outras.

Essa distância recebe este nome por ser proporcional à estatística qui-quadrado que testa a hipótese de independência entre linhas e colunas de uma tabela de contingência. Isso pode ser observado ao escrevermos a expressão da estatística  $\chi^2$ , dada por

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{\left( n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{n} \right)^2}{\frac{n_{i\cdot} n_{\cdot j}}{n}} \quad (2.3.9)$$

e ao multiplicar e dividir dentro do somatório por  $n_{i\cdot}^2$  e realizar operações adequadas, obtemos

$$\chi^2 = n \sum_{i=1}^I \frac{n_{i\bullet}}{n} \sum_{j=1}^J \left( \frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n} \right)^2 \frac{n}{n_{\bullet j}}. \quad (2.3.10)$$

Uma outra justificativa para se usar a distância qui-quadrado é que ela satisfaz a uma propriedade que é chamada de *Princípio da Equivalência Distribucional*. Este princípio estabelece que se dois perfis linha são idênticos (ou equivalentes distributivamente), então essas duas linhas da tabela de contingência podem ser somadas para gerar uma única linha. Esta nova linha terá o mesmo perfil e a massa igual à soma das massas das linhas somadas, sem que isso altere a distância entre os perfis coluna. Geometricamente, isso significa que se podem unir dois pontos de uma nuvem, que ocupam a mesma posição no “simplex”, sem afetar a geometria da outra nuvem. Esse resultado é particular à geometria da análise de correspondência. Lembrando que, quando unimos dois pontos de uma nuvem que ocupam posições distintas no “simplex”, as distâncias dos pontos da outra nuvem são alteradas. Esta propriedade é a que nos permite agrupar as linhas ou colunas de  $N$  que são linearmente dependentes.

Assim, definimos a distância para a nuvem dos perfis linha e -antes de definir o ajuste do subespaço- definiremos o quantificador que nos dá o quão a nuvem de pontos está dispersa no espaço  $\mathbf{R}^J$  e a transformação dos pontos para facilitar a compreensão da representação geométrica da nuvem de pontos.

### 2.3.2.1 A INÉRCIA

De acordo com a natureza dos dados a analisar, a nuvem terá maior ou menor dispersão em relação ao centróide. A dispersão poderá ser calculada como a média dos desvios dos vários pontos ao centróide. É natural medir o desvio entre dois pontos como a distância entre eles. Um quantificador da dispersão dos pontos dessa nuvem no espaço  $\mathbf{R}^J$  é chamado de inércia total dos perfis linha. Porém, antes de tratarmos disto, façamos um paralelo com os conceitos da física.

A noção de inércia, na mecânica, desempenha para os movimentos circulares o mesmo papel que a massa para os movimentos retilíneos. Um corpo apresenta tanta inércia em torno de um eixo quanta energia for necessária para colocá-lo em rotação em torno deste eixo. De maneira

mais precisa, a inércia elementar de um ponto que tem associada uma massa, em relação a um ponto qualquer, é o produto dessa massa pelo quadrado da distância entre os pontos.

Agora considerando a nuvem dos perfis linha como um sistema de massas pontuais, podemos calcular a inércia elementar de cada um dos perfis linha, em relação ao centróide. Assim, definimos a *inércia do i-ésimo perfil linha com relação ao centróide de R*, como o produto da massa  $r_i$  pelo quadrado da distância entre  $R_i$  e  $\bar{R}$ , dado por

$$In_{\bar{R}}(R_i) = r_i d^2(R_i, \bar{R}) \quad (2.3.11)$$

Caracterizamos, então, a dispersão da nuvem  $R$ , em relação ao centróide, como a soma das inércias de cada ponto em relação ao centróide, dada por

$$In_{\bar{R}}(R) = \sum_{i=1}^I In_{\bar{R}}(R_i) \quad (2.3.12)$$

e denominada *inércia total dos perfis linha em relação ao centróide de R*.

A dispersão de uma nuvem de pontos pode ser caracterizada não só em relação ao desvio do centróide, mas em relação a um ponto qualquer do mesmo espaço da nuvem. Assim podemos definir a *inércia do i-ésimo perfil linha em relação a um ponto qualquer  $y \in R^J$* , como o produto da massa  $r_i$  pelo quadrado da distância entre  $R_i$  e  $y$ , dado por

$$In_y(R_i) = r_i d^2(R_i, y). \quad (2.3.13)$$

Portanto, a dispersão da nuvem  $R$  em relação a  $y$  é dada por

$$In_y(R) = \sum_{i=1}^I In_y(R_i) \quad (2.3.14)$$

e denominada *inércia total dos perfis linha em relação a um ponto qualquer  $y \in R^J$* .

A inércia total da nuvem  $R$  em relação à  $\bar{R}$  e a inércia em relação a um ponto qualquer  $y \in R^J$  estão relacionadas pelo seguinte resultado que diz: a inércia total da nuvem  $R$  em relação a um ponto qualquer  $y \in R^J$  é igual à inércia total de  $R$  em relação ao centróide,  $\bar{R}$ , mais a inércia do ponto  $\bar{R}$  em relação ao ponto  $y$ , dado por

$$In_y(R) = In_{\bar{R}}(R) + In_y(\bar{R}). \quad (2.3.15)$$

Esse resultado é, às vezes, denominado Teorema de Huyghens, sua demonstração é simples e a faremos agora. Usando a definição de inércia dada em (2.3.11) e (2.3.12) e a definição de distância, em termos dos elementos de  $N$ , dada em (2.3.8) temos que a inércia total da nuvem  $R$ , em relação ao centróide, pode ser expressa por

$$In_{\bar{R}}(\mathbf{R}) = \sum_{i=1}^I r_i \left[ \sum_{j=1}^J \left( \frac{n_{ij}}{n_{i\bullet}} - \frac{n_{\bullet j}}{n} \right)^2 \frac{n}{n_{\bullet j}} \right].$$

Para simplificar a notação, vamos reescrever a expressão acima, em termos dos elementos da matriz  $\mathbf{R}$  e do vetor  $\mathbf{c}$ , onde representamos os elementos de  $\mathbf{R}$  por  $R_{ij}$ . Desta forma a expressão acima é dada por

$$In_{\bar{R}}(\mathbf{R}) = \sum_{i=1}^I r_i \left[ \sum_{j=1}^J (R_{ij} - c_j)^2 \frac{1}{c_j} \right].$$

Com esta notação, vamos demonstrar a expressão (2.3.15) de forma direta.

$$\begin{aligned} In_y(\mathbf{R}) &= \sum_{i=1}^I r_i \left[ \sum_{j=1}^J (R_{ij} - y_j)^2 \frac{1}{c_j} \right] \\ &= \sum_{j=1}^J \frac{1}{c_j} \left[ \sum_{i=1}^I r_i (R_{ij} - y_j)^2 \right] \\ &= \sum_{j=1}^J \frac{1}{c_j} \left[ \sum_{i=1}^I r_i \left( (R_{ij} - c_j)^2 + (c_j - y_j)^2 + 2(R_{ij} - c_j)(c_j - y_j) \right) \right] \\ &= \sum_{j=1}^J \frac{1}{c_j} \left[ \sum_{i=1}^I r_i (R_{ij} - c_j)^2 + \sum_{i=1}^I r_i (c_j - y_j)^2 + 2 \left( \sum_{i=1}^I r_i R_{ij} - c_j \right) (c_j - y_j) \right] \\ &= \sum_{i=1}^I r_i \sum_{j=1}^J (R_{ij} - c_j)^2 \frac{1}{c_j} + \sum_{i=1}^I r_i \sum_{j=1}^J (c_j - y_j)^2 \frac{1}{c_j} \\ &= \sum_{i=1}^I r_i \sum_{j=1}^J (R_{ij} - c_j)^2 \frac{1}{c_j} + \bar{r} \sum_{j=1}^J (c_j - y_j)^2 \frac{1}{c_j} \\ &= In_{\bar{R}}(\mathbf{R}) + In_y(\bar{\mathbf{R}}). \end{aligned}$$

A inércia total dos perfis linha em relação ao centróide é definida, por alguns autores, como a soma do quadrado das distâncias dos perfis linha ao centróide, ponderada pelas respectivas massas. Ela é obtida substituindo (2.3.11) em (2.3.12), dada por

$$In_{\bar{R}}(\mathbf{R}) = \sum_{i=1}^I r_i d^2(\mathbf{R}_i, \bar{\mathbf{R}}). \quad (2.3.16)$$

Substituindo a expressão da distância (2.3.8) e a das massas (2.3.2), dadas em termos dos elementos de  $\mathbf{N}$ , obtemos uma nova expressão da inércia total de  $\mathbf{R}$ . Comparando-a com a

expressão da estatística Qui-quadrado, dada em (2.3.10), obtemos a relação entre esses dois conceitos, relação esta dada por

$$In_{\bar{\mathbf{R}}}(\mathbf{R}) = \frac{\chi^2}{n}. \quad (2.3.17)$$

Podemos obter, ainda, uma nova expressão para a inércia total de  $\mathbf{R}$ , substituindo a distância dada em (2.3.8) em termos vetoriais, na expressão da inércia total de  $\mathbf{R}$  dada em (2.3.16), que pode ser expressa em termos matriciais por

$$In_{\bar{\mathbf{R}}}(\mathbf{R}) = \text{tr}[D_r (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') D_{\bar{\mathbf{R}}}^{-1} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}})'] \quad (2.3.18)$$

onde  $D_r$  é a matriz diagonal cuja diagonal é o vetor das massas.

A inércia total dada em (2.3.18) é o paralelo à variância total no caso contínuo e

$$\mathbf{S} = D_r (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') D_{\bar{\mathbf{R}}}^{-1} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}})' \quad (2.3.19)$$

é a matriz que tem como paralelo a matriz de variâncias e covariâncias. A matriz  $\mathbf{S}$  também pode ser vista como o produto interno ponderado dos vetores da matriz dos resíduos que é dada por

$$\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}' \quad (2.3.20)$$

e o produto interno é o definido pela métrica qui-quadrado.

Faremos, a seguir, um paralelo entre a Análise de Correspondência (AC) e a Análise de Componentes Principais (ACP). Observamos que enquanto nesta última temos  $P$  variáveis nas colunas, na primeira temos, nas colunas,  $J$  categorias de uma variável. Observamos também que, na ACP, temos a matriz dos produtos internos entre as variáveis - a chamada Matriz de Variâncias e Covariâncias- que apresenta, na diagonal principal, as variâncias entre as  $P$  variáveis. Por sua vez, na AC, quando trabalhando com os perfis linha, temos a matriz  $\mathbf{S}$  que nos dá os produtos internos entre as  $J$  categorias. Esta matriz, expressa em (2.3.19), apresenta na diagonal principal as inércias destas  $J$  categorias. Como aqui não faz sentido falar em co-inércias, a matriz  $\mathbf{S}$  será chamada apenas de Matriz de Inércias.

### 2.3.2.2 O ESPAÇO TRANSFORMADO

Como a ponderação da distância fica camuflada na representação geométrica, vamos, então, transformar os perfis mudando a distância euclidiana ponderada para distância euclidiana ordinária, tal que

$$\begin{aligned}
 d^2(\mathbf{R}_i, \mathbf{R}_j) &= (\mathbf{R}_i - \mathbf{R}_j)' \mathbf{D}_{\bar{\mathbf{R}}}^{-1} (\mathbf{R}_i - \mathbf{R}_j) \\
 &= (\mathbf{R}_i - \mathbf{R}_j)' \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} (\mathbf{R}_i - \mathbf{R}_j) \\
 &= (\mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \mathbf{R}_i - \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \mathbf{R}_j)' (\mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \mathbf{R}_i - \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \mathbf{R}_j)
 \end{aligned}$$

Assim cada eixo coordenado fica estendido na proporção de  $1/c_j$  e tendo, portanto, uma escala diferente, como podemos observar, na figura 2 abaixo, a representação do  $i$ -ésimo perfil linha e o seu correspondente perfil transformado, no espaço  $\mathbf{R}^3$ .

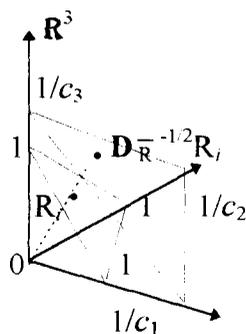


Figura 2: Representação do Perfil Linha transformado no Espaço  $\mathbf{R}^3$

Devido a essa transformação a matriz  $\mathbf{S}$  pode ser expressa por

$$\mathbf{S} = \mathbf{X}'\mathbf{X} \quad (2.3.21)$$

onde

$$\mathbf{X} = \mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \quad (2.3.22)$$

é a matriz dos resíduos -ponderados- dos perfis linha com relação ao centróide no espaço transformado, pois de (2.3.19) temos

$$\begin{aligned}
 \mathbf{S} &= (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}')' \mathbf{D}_r (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1} \\
 &= (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}')' \mathbf{D}_r^{1/2} \mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \\
 &= \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}')' \mathbf{D}_r^{1/2} \mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2} \\
 &= [\mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2}]' [\mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2}] \\
 &= [\mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2}]' [\mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1}_i \bar{\mathbf{R}}') \mathbf{D}_{\bar{\mathbf{R}}}^{-1/2}].
 \end{aligned}$$

### 2.3.3 O AJUSTE DO SUBESPAÇO

Até o momento definimos uma nuvem de pontos e suas massas, em um espaço estruturado por uma métrica euclidiana generalizada. Mas lembremos que a Análise de Correspondência Simples tem como objetivo mostrar as possíveis associações, heterogeneidade ou dependência entre as linhas e colunas da tabela de contingência, em um espaço vetorial de dimensão menor que a original. Assim, ajustar um subespaço de menor dimensão a uma nuvem de pontos significa rotacionar os eixos de modo que seja possível ver a nuvem de um ângulo diferente, que nos mostre as projeções dos pontos com a maior dispersão. Os novos eixos que definem esse subespaço indicam as direções de maior inércia dos dados, e são perpendiculares entre si.

Para a nuvem dos perfis linha, buscaremos inicialmente o subespaço que denominamos de *Subespaço Ajustado* e cuja dimensão é  $K = \min\{I, J\} - 1$ . Este subespaço nos dá o ajuste perfeito dos  $I$  pontos de  $\mathbf{R}$ , no sentido de que a soma -ponderada pela massa- do quadrado das distâncias dos pontos ao subespaço ajustado seja mínima. Em seguida, para fazer a análise, buscaremos o subespaço vetorial de dimensão  $K^* \leq 3$  onde são considerados apenas os primeiros eixos. Este subespaço é denominado *Subespaço Ótimo* e nos permite visualizar a representação dos pontos de  $\mathbf{R}$ . Quando afirmamos, acima, que o espaço ajustado tem dimensão  $K$ , estamos considerando que  $\mathbf{N}$  não tem nenhuma das linhas ou colunas linearmente dependente de outra. Caso isso ocorra, a dimensão do espaço ajustado será menor do que  $K$ .

Para obter estes subespaços, começaremos por buscar um subespaço unidimensional, ou seja, buscaremos uma linha que melhor se ajuste à nuvem de pontos por quadrados mínimos ponderados, usando as massas das linhas  $r_i$  como pesos dos pontos. O problema do ajuste por quadrados mínimos pode ser delineado de duas maneiras: encontrar um mínimo para o quadrado das distâncias dos pontos à reta ou determinar a distância máxima entre as projeções dos pontos sobre a reta e a origem. Depois, buscaremos um segundo subespaço unidimensional, que seja ortogonal ao primeiro, mas que tenha as mesmas características que o anterior e assim por diante. O problema é análogo ao de achar as componentes principais de um conjunto de  $I$  observações e  $J$  variáveis, com uma generalização simples no uso de uma métrica generalizada e na ponderação das observações.

Para resolver este problema, vamos começar por lembrar o que foi feito até agora. Primeiramente, definimos a matriz  $\mathbf{R}$ , em (2.3.3), que nos dá as coordenadas dos  $I$  pontos no espaço  $\mathbf{R}^J$  estruturados pela métrica qui-quadrado. Depois a matriz  $\mathbf{R}^{-1}\bar{\mathbf{R}}'$ , em (2.3.20), nos dá as coordenadas centralizadas desses pontos, no mesmo espaço. Ao fazer a transformação dos pontos, a matriz  $(\mathbf{R} - \mathbf{1}_I\bar{\mathbf{R}}') D_{\bar{\mathbf{R}}}^{-1/2} = D_r^{-1/2} \mathbf{X}$ , onde  $\mathbf{X}$  é dado em (2.3.22), nos dá as coordenadas centralizadas dos pontos, no espaço  $\mathbf{R}^J$ -transformado, isto é, a métrica foi a euclidiana simples, mas cada ponto tem uma massa a considerar. É nesse espaço transformado que vamos, a partir de agora, encontrar os vetores direcionais dos eixos principais.

Seja  $\mathbf{v} \in \mathbf{R}^J$  o vetor que queremos encontrar, unitário e cuja direção nos mostre a maior “variabilidade” dos pontos da nuvem, ou seja, queremos encontrar um vetor direcional  $\mathbf{v}$ , tal que as “projeções” dos pontos sobre esse vetor tenha distância máxima à origem. Como cada ponto tem uma massa, quando falamos em “variabilidade” e “projeção”, devemos considerá-la. Assim, queremos encontrar uma direção em que os pontos projetados tenham as maiores inércias.

Cada linha de  $D_r^{-1/2} \mathbf{X}$  é um ponto do espaço  $\mathbf{R}^J$ -transformado e o produto

$$D_r^{-1/2} \mathbf{X}\mathbf{v} \quad (2.3.23)$$

é um ponto com  $I$  componentes onde as componentes são os produtos escalares das linhas de  $D_r^{-1/2} \mathbf{X}$  por  $\mathbf{v}$ , que são as coordenadas das projeções dos  $I$  pontos da nuvem sobre o eixo  $\mathbf{v}$ .

Assim a inércia das projeções dos pontos da nuvem sobre  $\mathbf{v}$  é dada por

$$(D_r^{-1/2} \mathbf{X}\mathbf{v})' D_r(D_r^{-1/2} \mathbf{X}\mathbf{v}) = \mathbf{v}' \mathbf{X}' \mathbf{X} \mathbf{v} = \mathbf{v}' \mathbf{S} \mathbf{v} \quad (2.3.24)$$

que é a expressão a ser maximizada em  $\mathbf{v}$ , com a restrição de que  $\mathbf{v}$  seja unitário.

Vamos, então, encontrar vetores  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_J$ , tal que

$$\mathbf{v}_1' \mathbf{S} \mathbf{v}_1 \text{ seja máxima e } \mathbf{v}_1' \mathbf{v}_1 = 1$$

$$\mathbf{v}_2' \mathbf{S} \mathbf{v}_2 \text{ seja máxima e } \mathbf{v}_2' \mathbf{v}_2 = 1, \mathbf{v}_1' \mathbf{v}_2 = 0$$

$$\vdots$$

$$\mathbf{v}_J' \mathbf{S} \mathbf{v}_J \text{ seja máxima e } \mathbf{v}_J' \mathbf{v}_J = 1, \mathbf{v}_1' \mathbf{v}_2 = 0, \dots, \mathbf{v}_{J-1}' \mathbf{v}_J = 0$$

Isso é equivalente a encontrar uma base ortonormal do espaço  $\mathbf{R}^J$ , com a restrição de que as projeções dos perfis, em cada eixo, tenham inércia máxima.

Esse problema está resolvido, de forma mais geral e detalhada, em vários textos, por exemplo, Lebart, Morineau e Warwick (1984). Aqui utilizamos as ferramentas matemáticas necessárias, no contexto do espaço vetorial dos perfis linha.

Como esse problema é de maximização sujeita a uma restrição, utilizemos o método de multiplicadores de Lagrange. A equação de Lagrange para esse problema é

$$L(\mathbf{x}, \lambda) = \mathbf{x}'\mathbf{S}\mathbf{x} - \lambda_1(\mathbf{x}'\mathbf{x} - 1),$$

onde  $\lambda_1$  é o multiplicador de Lagrange. Como  $\mathbf{x}'\mathbf{S}\mathbf{x}$  e  $\mathbf{x}'\mathbf{x}$  são formas quadráticas, temos que

$$\nabla[\mathbf{x}'\mathbf{S}\mathbf{x}] = 2\mathbf{S}\mathbf{x}$$

e

$$\nabla[\mathbf{x}'\mathbf{x}] = 2\mathbf{x}.$$

Portanto, o gradiente da equação de Lagrange é dado por

$$\nabla[L(\mathbf{x}, \lambda_1)] = 2\mathbf{S}\mathbf{x} - 2\lambda_1\mathbf{x}$$

assim, obtemos o ponto crítico quando igualamos as derivadas a zero, isto é,

$$\nabla[L(\mathbf{v}_1, \lambda_1)] = 2\mathbf{S}\mathbf{v}_1 - 2\lambda_1\mathbf{v}_1 = \mathbf{0}.$$

Daí obtemos

$$\mathbf{S}\mathbf{v}_1 = \lambda_1\mathbf{v}_1. \quad (2.3.25)$$

Portanto temos que  $\mathbf{v}_1$ , o vetor que faz com que  $\mathbf{v}_1'\mathbf{S}\mathbf{v}_1$  tenha o valor máximo, é o autovetor da matriz  $\mathbf{S}$  associado ao autovalor  $\lambda_1$ . Quando pré-multiplicamos a equação (2.3.25) por  $\mathbf{v}_1'$  e levando em consideração que  $\mathbf{v}_1'\mathbf{v}_1 = 1$ , temos que

$$\mathbf{v}_1'\mathbf{S}\mathbf{v}_1 = \lambda_1 \quad (2.3.26)$$

portanto  $\lambda_1$  será o maior autovalor de  $\mathbf{S}$ .

Procuremos, então, o vetor  $\mathbf{v}_2$  que maximize a forma quadrática  $\mathbf{v}_2'\mathbf{S}\mathbf{v}_2$  tal que seja ortogonal a  $\mathbf{v}_1$  ( $\mathbf{v}_2'\mathbf{v}_1 = 0$ ) e unitário ( $\mathbf{v}_2'\mathbf{v}_2 = 1$ ). Analogamente, aplicando o método de Lagrange, obtemos a seguinte equação

$$L(\mathbf{x}, \lambda_2, \eta) = \mathbf{x}'\mathbf{S}\mathbf{x} - \lambda_2(\mathbf{x}'\mathbf{x} - 1) - \eta(\mathbf{x}'\mathbf{v}_1)$$

onde  $\lambda_2$  e  $\eta$  são os multiplicadores de Lagrange. A solução ótima é encontrada quando igualamos as derivadas da equação a zero,

$$\nabla[L(\mathbf{v}_2, \lambda_2, \eta)] = 2\mathbf{S}\mathbf{v}_2 - 2\lambda_2\mathbf{v}_2 - 2\eta\mathbf{v}_1 = \mathbf{0}.$$

Se pré-multiplicarmos a equação acima por  $\mathbf{v}_2'$ , levando em consideração  $\mathbf{v}_2'\mathbf{v}_1 = 0$  e  $\mathbf{v}_2'\mathbf{v}_2 = 1$ , obteremos

$$\mathbf{v}_2'\mathbf{S}\mathbf{v}_2 = \lambda_2. \quad (2.3.27)$$

Se pós-multiplicarmos a mesma equação por  $\mathbf{v}_1'$ , vemos que  $\eta = 0$  e se escrevermos

$$\mathbf{v}_1'\mathbf{S}\mathbf{v}_2 = \lambda_2\mathbf{v}_1'\mathbf{v}_2 = 0$$

teremos

$$\mathbf{S} \mathbf{v}_2 = \lambda_2 \mathbf{v}_2 \quad (2.3.28)$$

e portanto  $\mathbf{v}_2$  será autovetor de  $\mathbf{S}$  e estará associado ao seu segundo maior autovalor. Esta prova é extensível ao caso do  $k$ -ésimo vetor  $\mathbf{v}_k$  ( $k=1, \dots, J$ ), que maximize a forma quadrática  $\mathbf{v}_k' \mathbf{S} \mathbf{v}_k$  com  $\mathbf{v}_k' \mathbf{v}_k = 1$  e  $\mathbf{v}_k' \mathbf{v}_{k-1} = 0$ ,  $\mathbf{v}_k' \mathbf{v}_{k-2} = 0$ , ...,  $\mathbf{v}_k' \mathbf{v}_1 = 0$ . Então obteremos

$$\mathbf{v}_k' \mathbf{S} \mathbf{v}_k = \lambda_k \quad (2.3.29)$$

e

$$\mathbf{S} \mathbf{v}_k = \lambda_k \mathbf{v}_k \quad (2.3.30)$$

Assim, vemos que a solução do problema está nos autovalores e autovetores de  $\mathbf{S}$  e, portanto, podemos dizer que  $\mathbf{S}$  é a matriz que define o ajuste do subespaço procurado.

Escrevendo este resultado em forma de matrizes temos

$$\mathbf{S} \mathbf{V} = \mathbf{V} \mathbf{D}_\lambda \quad \text{tal que} \quad \mathbf{V}' \mathbf{V} = \mathbf{I} \quad (2.3.31)$$

onde  $\mathbf{V} = [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_J]$  e  $\mathbf{D}_\lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_J \end{bmatrix}$

o que implica

$$\mathbf{S} = \mathbf{V} \mathbf{D}_\lambda \mathbf{V}' \quad \text{com} \quad \mathbf{V}' \mathbf{V} = \mathbf{I} \quad (2.3.32)$$

que é a chamada *Decomposição Espectral* de  $\mathbf{S}$ . Assim os vetores direcionais procurados são exatamente os autovetores da matriz  $\mathbf{S}$  e cada um destes vetores (eixos) tem associado uma inércia dada pelo autovalor  $\lambda_k$ , que indica o quanto a nuvem está dispersa nessa direção. Desta forma, podemos dizer que  $\mathbf{S}$  é a matriz que define o ajuste do subespaço.

As coordenadas das projeções dos pontos da nuvem dos perfis linha nos eixos considerados são dadas como expresso em (2.3.23) pela matriz

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{X} \mathbf{V} \quad (2.3.33)$$

que é chamada *Matriz das Coordenadas Principais* dos perfis linha, onde  $\mathbf{V}$  é a matriz dos autovetores de  $\mathbf{S}$ .

Haverá situações em que a interpretação será facilitada se cada eixo tiver inércia igual a 1 (um). Isso é feito dividindo as componentes da inércia dos pontos em cada eixo por sua inércia principal. As coordenadas nestes eixos são denominadas de coordenadas padronizadas e são dadas pela seguinte matriz

$$\Phi = F D_{\lambda}^{-1/2} \quad (2.3.34)$$

que é chamada *matriz das coordenadas padronizadas* dos perfis linha. Na seção 2.6.1, ao definirmos a contribuição absoluta, justificamos esta definição.

Veremos, na próxima seção, algumas relações que nos permitem simplificar a expressão das coordenadas principais e padronizadas.

### 2.3.3.1 ESPAÇO VETORIAL NÃO CENTRALIZADO

O ajuste do subespaço ótimo foi feito em um espaço vetorial onde a origem do sistema era o centróide da nuvem, isto é, em um espaço centralizado. Assim, a matriz a ser diagonalizada é a matriz  $S = X'X$  onde  $X$  é a matriz dos perfis centralizados, ponderados e transformados, dada em (2.3.22). Com o intuito de simplificar a expressão das coordenadas principais vamos mostrar algumas relações entre o espaço vetorial centralizado e o não centralizado. Para isso definiremos a matriz dos perfis ponderados, transformados porém não centralizados, dada por

$$X^* = D_r^{1/2} R D_{\bar{R}}^{-1/2} \quad (2.3.35)$$

Assim, a matriz a ser diagonalizada é então dada por

$$S^* = X^{*'} X^* \quad (2.3.36)$$

A seguir, veremos duas relações entre estes dois espaços.

$$1) \text{ O vetor } v_a = \left( \sqrt{\frac{n_{\cdot 1}}{n}} \quad \cdots \quad \sqrt{\frac{n_{\cdot j}}{n}} \right)' = D_{\bar{R}}^{-1/2} \mathbf{1} \text{ é autovetor de } S \text{ associado ao autovalor}$$

0 e é autovetor de  $S^*$  associado ao autovalor 1, pois

$$\begin{aligned} i) X v_a &= [D_r^{1/2} (R - \mathbf{1} \bar{R}') D_{\bar{R}}^{-1/2}] v_a \\ &= [D_r^{1/2} (R - \mathbf{1} \bar{R}') D_{\bar{R}}^{-1/2}] D_{\bar{R}}^{-1/2} \mathbf{1} \\ &= [D_r^{1/2} (R - \mathbf{1} \bar{R}') D_{\bar{R}}^{-1/2} D_{\bar{R}}^{-1/2} \mathbf{1}] \\ &= [D_r^{1/2} (R - \mathbf{1} \bar{R}') \mathbf{1}] \\ &= [D_r^{1/2} (R \mathbf{1} - \mathbf{1} \bar{R}' \mathbf{1})] \\ &= [D_r^{1/2} (\mathbf{1} - \mathbf{1})] \\ &= \mathbf{0}. \end{aligned}$$

Portanto  $\mathbf{v}_a' \mathbf{S} \mathbf{v}_a = \mathbf{v}_a' \mathbf{X}' \mathbf{X} \mathbf{v}_a = (\mathbf{X} \mathbf{v}_a)' (\mathbf{X} \mathbf{v}_a) = \mathbf{0}' \mathbf{0} = 0$ , implicando

$$\mathbf{S} \mathbf{v}_a = 0 \mathbf{v}_a \quad (2.3.37)$$

$$\begin{aligned} ii) \mathbf{X}^* \mathbf{v}_a &= [\mathbf{D}_r^{1/2} \mathbf{R} \mathbf{D}_R^{-1/2}] \mathbf{v}_a \\ &= \mathbf{D}_r^{1/2} \mathbf{R} \mathbf{D}_R^{-1/2} \mathbf{D}_R^{-1/2} \mathbf{1} \\ &= \mathbf{D}_r^{1/2} \mathbf{R} \mathbf{1} \\ &= \mathbf{D}_r^{1/2} \mathbf{1} \\ &= \mathbf{v}_x. \end{aligned}$$

Logo,  $\mathbf{v}_a' \mathbf{S}^* \mathbf{v}_a = \mathbf{v}_a' \mathbf{X}^{**} \mathbf{X}^* \mathbf{v}_a = (\mathbf{X}^* \mathbf{v}_a)' (\mathbf{X}^* \mathbf{v}_a) = \mathbf{v}_x' \mathbf{v}_x = 1$  concluindo-se que

$$\mathbf{S}^* \mathbf{v}_a = 1 \mathbf{v}_a \quad (2.3.38)$$

É óbvio, portanto, que os autovalores zero ou um dados em (2.3.37) e (2.3.38), conforme os dados tenham sido ou não centralizados, são triviais pois associam-se a um autovetor que é a diagonal do espaço ou transformação linear dele. A projeção sobre esta direção é nula, no caso centralizado, e um, no caso não centralizado.

2) Todos os autovetores de  $\mathbf{S}$  distintos de  $\mathbf{v}_a$  são também autovetores de  $\mathbf{S}^*$  associados aos mesmos autovalores.

Pois, usando o fato de que  $\mathbf{v}_a$  é ortogonal a todos os autovetores de  $\mathbf{S}$ , isto é,  $\mathbf{v}_a' \mathbf{v}_{j-1} = 0$ ,  $\mathbf{v}_a' \mathbf{v}_{j-2} = 0, \dots, \mathbf{v}_a' \mathbf{v}_1 = 0$  temos, que

$$\begin{aligned} \mathbf{X} \mathbf{v}_k &= [\mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1} \bar{\mathbf{R}}') \mathbf{D}_R^{-1/2}] \mathbf{v}_k \\ &= [\mathbf{D}_r^{1/2} (\mathbf{R} - \mathbf{1} \bar{\mathbf{R}}') \mathbf{D}_R^{-1/2} \mathbf{v}_k] \\ &= [\mathbf{D}_r^{1/2} (\mathbf{R} \mathbf{D}_R^{-1/2} \mathbf{v}_k - \mathbf{1} \bar{\mathbf{R}}' \mathbf{D}_R^{-1/2} \mathbf{v}_k)] \\ &= [\mathbf{D}_r^{1/2} (\mathbf{R} \mathbf{D}_R^{-1/2} \mathbf{v}_k - \mathbf{1} (\mathbf{D}_R^{-1/2} \mathbf{1})' \mathbf{v}_k)] \\ &= [\mathbf{D}_r^{1/2} (\mathbf{R} \mathbf{D}_R^{-1/2} \mathbf{v}_k - \mathbf{1} (\mathbf{v}_a)' \mathbf{v}_k)] \\ &= \mathbf{D}_r^{1/2} \mathbf{R} \mathbf{D}_R^{-1/2} \mathbf{v}_k \\ &= \mathbf{X}^* \mathbf{v}_k \end{aligned}$$

e portanto  $\mathbf{v}_k' \mathbf{S} \mathbf{v}_k = \mathbf{v}_k' \mathbf{X}' \mathbf{X} \mathbf{v}_k = \mathbf{v}_k' \mathbf{X}^{**} \mathbf{X}^* \mathbf{v}_k = \mathbf{v}_k' \mathbf{S}^* \mathbf{v}_k \Rightarrow \mathbf{S} \mathbf{v}_k = \mathbf{S}^* \mathbf{v}_k = \lambda_k \mathbf{v}_k$  de forma matricial  $\mathbf{S} \mathbf{V} = \mathbf{S}^* \mathbf{V}$  e

$$\mathbf{X} \mathbf{V} = \mathbf{X}^* \mathbf{V}. \quad (2.3.39)$$

Vemos então que a decomposição da matriz  $\mathbf{S}^*$ , contém a decomposição da matriz  $\mathbf{S}$ , isto é,

$$\mathbf{S}^* \mathbf{V} = \mathbf{V} \mathbf{D}_\lambda \text{ tal que } \mathbf{V}' \mathbf{V} = \mathbf{I}$$

$$\text{onde } \mathbf{V} = [ \mathbf{v}_a \mid \mathbf{v}_1 \mid \dots \mid \mathbf{v}_{j-1} ] \text{ e } \mathbf{D}_\lambda = \begin{bmatrix} 1 & 0 & & 0 \\ 0 & \lambda_1 & & 0 \\ & & \ddots & \\ 0 & 0 & & \lambda_{j-1} \end{bmatrix}$$

e

$$\mathbf{S}\mathbf{V} = \mathbf{V}\mathbf{D}_\lambda \text{ tal que } \mathbf{V}'\mathbf{V} = \mathbf{I}$$

$$\text{onde } \mathbf{V} = [ \mathbf{v}_1 \mid \dots \mid \mathbf{v}_{j-1} \mid \mathbf{v}_a ] \text{ e } \mathbf{D}_\lambda = \begin{bmatrix} \lambda_1 & & 0 & 0 \\ & \ddots & & \\ 0 & & \lambda_{j-1} & 0 \\ 0 & & 0 & 0 \end{bmatrix}$$

Vemos, portanto, que podemos trabalhar tanto no espaço centralizado como no espaço não centralizado, desde que consideremos apenas os autovetores não triviais, ou seja, no espaço centralizado desprezamos o autovetor associado ao autovalor zero (0) e no espaço não centralizado desprezamos o autovetor associado ao autovalor 1 (um).

A partir da relação (2.3.39) podemos reescrever as coordenadas principais em termos dos perfis. Ao substituí-la em  $\mathbf{F}$  dado em (2.3.33), obtemos

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{X}^* \mathbf{V}. \quad (2.3.40)$$

Substituindo, na igualdade acima, o  $\mathbf{X}^*$  dado em (2.3.35), temos uma outra expressão das coordenadas principais em termos dos perfis dada por

$$\mathbf{F} = \mathbf{R} \mathbf{D}_R^{-1/2} \mathbf{V}. \quad (2.3.41)$$

Assim o subespaço  $K$ -dimensional, que melhor se ajusta aos dados em termos dos quadrados mínimos ponderados, está definido pelos  $K$  autovetores  $\mathbf{v}_k$ ,  $k=1, \dots, K$ , correspondentes aos  $K$  maiores autovalores  $\lambda_k$ ,  $k=1, \dots, K$ , de  $\mathbf{S}$ . Para não sobrecarregar em notações, expressamos as coordenadas dos subespaços ajustado e ótimo através da expressão dada em (2.3.33) e (2.3.41), sendo as dimensões das matrizes definidas pelo contexto.

Estudadas as relações entre os pontos-linha, dediquemo-nos ao estudo das relações entre os pontos coluna.

## 2.4 RELAÇÕES ENTRE AS COLUNAS

O estudo que será apresentado nesta seção é o dual do realizado na seção anterior. Estaremos procedendo de forma análoga e simétrica ao caso das relações entre linhas no espaço  $\mathbf{R}^I$ . Contudo, agora, estaremos representando as *colunas* da matriz de dados no espaço dos perfis coluna para obter uma nuvem de  $J$  pontos no espaço  $\mathbf{R}^J$ . A analogia é total e poderíamos omitir esta seção, mantida apenas por razão das notações que serão necessárias posteriormente.

### 2.4.1 A NUVEM DE PONTOS NO ESPAÇO $\mathbf{R}^J$

Os pontos, que representam as colunas da matriz  $\mathbf{N}$ , formam uma nuvem de  $J$  pontos no espaço  $\mathbf{R}^J$  onde cada ponto são os perfis das colunas ao longo das linhas. Dado por

$$C_j = \left( \frac{n_{1j}}{n_{\bullet j}}, \dots, \frac{n_{Ij}}{n_{\bullet j}} \right), j = 1, \dots, J \quad (2.4.1)$$

que é denominado o  $j$ -ésimo *perfil coluna* de  $\mathbf{N}$  onde  $n_{\bullet j} \neq 0$  conforme restrição em (2.2.3) e obs.:1, na página 11. As massas associadas a esses perfis são dadas por

$$c_j = \frac{n_{\bullet j}}{n}, j = 1, \dots, J. \quad (2.4.2)$$

A matriz que nos dá a nuvem dos  $J$  perfis coluna no espaço  $\mathbf{R}^J$  é dada por

$$\mathbf{C} = \begin{bmatrix} \frac{n_{11}}{n_{\bullet 1}} & \dots & \frac{n_{1J}}{n_{\bullet J}} \\ \frac{n_{\bullet 1}}{n} & \ddots & \frac{n_{\bullet J}}{n} \\ \frac{n_{I1}}{n_{\bullet 1}} & \dots & \frac{n_{IJ}}{n_{\bullet J}} \\ \frac{n_{\bullet 1}}{n} & \dots & \frac{n_{\bullet J}}{n} \end{bmatrix} = [C_1' | \dots | C_J']' = \mathbf{D}_c^{-1} \mathbf{P}' \quad (2.4.3)$$

e denominada *matriz dos perfis coluna*. O *vetor de massas* associado à nuvem  $\mathbf{C}$ , dos perfis coluna, que nos dá as frequências relativas das colunas de  $\mathbf{N}$  é dado por

$$\mathbf{c} = \left( \frac{n_{\bullet 1}}{n} \quad \dots \quad \frac{n_{\bullet J}}{n} \right), \quad (2.4.4)$$

Portanto as massas dos perfis são as marginais  $c_j$ , dadas pelo vetor  $\mathbf{c}$  em (2.2.6).

O *centróide* dos perfis coluna, o ponto que nos mostra o perfil da nuvem em relação às linhas, é denotado por  $\bar{C}$  e dado por

$$\bar{C} = \left( \frac{n_{1\bullet}}{n} \quad \dots \quad \frac{n_{I\bullet}}{n} \right), \quad (2.4.5)$$

Sua massa é dada por

$$\bar{c} = \sum_{j=1}^J c_j. \quad (2.4.6)$$

Aqui também temos que o centróide é o perfil médio da nuvem, isto é, é a média dos perfis coluna ponderada pelas suas respectivas massas

$$\bar{C} = \frac{1}{\bar{c}} \sum_{j=1}^J c_j C_j = C^* c. \quad (2.4.7)$$

Como em tabelas de contingência simples (como a tabela 1) temos que  $\bar{c}=1$ , decorre que o centróide é igual à soma dos pontos  $C_j, j=1, \dots, J$ , ponderada pelas suas respectivas massas  $c_j$ .

### 2.4.2 A MÉTRICA DA NUVEM NO ESPAÇO $\mathbf{R}^I$

Após definirmos a nuvem de pontos em (2.4.1), vamos agora estruturar o espaço dos perfis coluna  $\mathbf{R}^I$ , com a distância qui-quadrado que é definida tomando como base o centróide da nuvem. Assim a *distância qui-quadrado* para a nuvem dos perfis coluna é expressa por

$$d^2(C_j, \bar{C}) = \sum_{i=1}^I \left( \frac{n_{ij}}{n_{i\bullet}} - \frac{n_{i\bullet}}{n} \right)^2 \frac{n}{n_{i\bullet}} = (C_j - \bar{C})' D_{\bar{c}}^{-1} (C_j - \bar{C}) \quad (2.4.8)$$

onde  $D_{\bar{c}}^{-1}$  é a matriz diagonal em que na diagonal temos o inverso dos elementos do centróide.

Aqui o *Princípio da Equivalência Distribucional* estabelece que se dois perfis coluna são idênticos (ou equivalentes distributivamente), então essas duas colunas da tabela de contingência podem ser somadas para gerar uma única coluna. Esta nova coluna terá o mesmo perfil e a massa igual à soma das massas das colunas somadas, sem que isso altere a distância entre os perfis linha. Geometricamente, isso significa que se podem unir dois pontos da nuvem dos perfis coluna, que ocupam a mesma posição no “simplex”, sem alterar as distâncias dos pontos da nuvem dos perfis linha.

Assim definimos a distância para a nuvem de pontos dos perfis coluna e antes de definir o ajuste do subespaço, definiremos o quantificador, que nos dá o quão a nuvem de pontos está dispersa no espaço  $\mathbf{R}^l$  e a transformação dos pontos, para compreensão da representação geométrica.

### 2.4.2.1 INÉRCIA TOTAL E ESPAÇO TRANSFORMADO

Da mesma forma que na seção (2.3.2.1), considerando a nuvem dos perfis coluna como um sistema de massas pontuais, vamos calcular a inércia elementar de cada um dos perfis coluna, em relação ao centróide. Assim, a *inércia do j-ésimo perfil coluna com relação ao centróide de C*, é o produto da massa  $c_j$  pelo quadrado da distância entre  $C_j$  e  $\bar{C}$ , dada por

$$In_{\bar{C}}(C_j) = c_j d^2(C_j, \bar{C}). \quad (2.4.9)$$

Portanto, a dispersão da nuvem  $\mathbf{C}$ , em relação ao centróide, é definida como a soma das inércias de cada ponto em relação ao centróide, dada por

$$In_{\bar{C}}(\mathbf{C}) = \sum_{j=1}^J In_{\bar{C}}(C_j) \quad (2.4.10)$$

e denominada *inércia total dos perfis coluna em relação ao centróide de C*.

A dispersão de uma nuvem de pontos pode ser caracterizada não só em relação ao desvio do centróide, mas em relação a um ponto qualquer do mesmo espaço da nuvem. Assim a *inércia do j-ésimo perfil coluna em relação a um ponto qualquer  $y \in \mathbf{R}^l$* , é dado por

$$In_y(C_j) = c_j d^2(C_j, y). \quad (2.4.11)$$

Portanto, a dispersão da nuvem  $\mathbf{C}$  em relação a  $y$  é dada por

$$In_y(\mathbf{C}) = \sum_{j=1}^J In_y(C_j) \quad (2.4.12)$$

e denominada *inércia total dos perfis coluna em relação a um ponto qualquer  $y \in \mathbf{R}^l$* .

Da mesma forma que para a nuvem dos perfis linha, temos que a inércia total da nuvem dos perfis coluna  $\mathbf{C}$ , em relação a um ponto qualquer  $y \in \mathbf{R}^l$  é igual à inércia total da nuvem  $\mathbf{C}$  em relação ao centróide  $\bar{C}$ , mais a inércia do ponto  $\bar{C}$  em relação ao ponto  $y$  dado por

$$In_y(\mathbf{C}) = In_{\bar{C}}(\mathbf{C}) + In_y(\bar{C}). \quad (2.4.13)$$

A inércia total dos perfis coluna em relação ao centróide é apresentada, às vezes, como a soma do quadrado das distâncias dos perfis linha ao centróide, ponderada pelas respectivas massas, dada por

$$In_{\bar{c}}(\mathbf{C}) = \sum_{j=1}^J c_j d^2(C_j, \bar{C}). \quad (2.4.14)$$

A relação entre a estatística Qui-quadrado e a inércia total de  $\mathbf{C}$ , é dada por

$$In_{\bar{c}}(\mathbf{C}) = \frac{\chi^2}{n}. \quad (2.4.15)$$

Uma nova expressão para a inércia total de  $\mathbf{C}$  é dada em termos matriciais por

$$In_{\bar{c}}(\mathbf{C}) = \text{tr}[\mathbf{D}_c (\mathbf{C} - \mathbf{1}_J \bar{C}')' \mathbf{D}_{\bar{c}}^{-1} (\mathbf{C} - \mathbf{1}_J \bar{C}')'] \quad (2.4.16)$$

onde  $\mathbf{D}_c$  é a matriz diagonal cuja diagonal é o vetor das massas.

A inércia total dada em (2.4.16) é o paralelo à variância total no caso contínuo e

$$\mathbf{Z} = \mathbf{D}_c (\mathbf{C} - \mathbf{1}_J \bar{C}')' \mathbf{D}_{\bar{c}}^{-1} (\mathbf{C} - \mathbf{1}_J \bar{C}')' \quad (2.4.17)$$

é a matriz que tem como paralelo a matriz de variâncias e covariâncias.

A distância no espaço transformado é dada por

$$d^2(C_i, C_j) = (\mathbf{D}_{\bar{c}}^{-1/2} C_i - \mathbf{D}_{\bar{c}}^{-1/2} C_j)' (\mathbf{D}_{\bar{c}}^{-1/2} C_i - \mathbf{D}_{\bar{c}}^{-1/2} C_j)$$

e, portanto, com essa transformação  $\mathbf{Z}$  também pode ser expressa por

$$\mathbf{Z} = \mathbf{Y}'\mathbf{Y} \quad (2.4.18)$$

onde

$$\mathbf{Y} = \mathbf{D}_c^{1/2} (\mathbf{C} - \mathbf{1}_J \bar{C}')' \mathbf{D}_{\bar{c}}^{-1/2} \quad (2.4.19)$$

é a matriz dos resíduos, ponderados, dos perfis coluna com relação ao centróide no espaço transformado.

### 2.4.3 O AJUSTE DO SUBESPAÇO

Como no caso das linhas, queremos encontrar vetores  $u_1, \dots, u_l$ , que nos dão uma base ortonormal do espaço vetorial  $\mathbb{R}^l$ . Sendo que os  $K$  primeiros eixos definem os eixos principais do subespaço ajustado. Vimos na seção 2.3.3 que isso é dado pelos autovetores de uma matriz simétrica, que aqui é dada pela matriz  $Z$ , dada em (2.4.17), quando maximizada a forma quadrática

$$u'Zu$$

tal que  $u$  seja unitário. Os vetores  $u_1, \dots, u_l$ , são dados pela decomposição espectral de  $Z$  que é expressa por

$$ZU = UD_\eta \text{ com } U'U = I \quad (2.4.20)$$

ou

$$Z = UD_\eta U' \text{ com } U'U = I \quad (2.4.21)$$

onde  $U = [u_1 | \dots | u_l]$  é a matriz dos autovetores de  $Z$  e  $D_\eta = \begin{bmatrix} \eta_1 & & 0 \\ & \ddots & \\ 0 & & \eta_l \end{bmatrix}$  é a matriz diagonal

dos autovalores associado aos autovetores de  $Z$ . Assim,  $Z$  é a matriz que nos dá o ajuste do subespaço  $\mathbb{R}^l$ .

Desta forma, a *matriz das coordenadas principais* dos perfis coluna é, então, dada por

$$G = D_c^{-1/2} YU \quad (2.4.22)$$

e a *matriz das coordenadas padronizadas* dos perfis linha é dada por

$$\Gamma = G D_\eta^{-1/2} \quad (2.4.23)$$

Veremos, na próxima subseção, algumas relações que nos permitem simplificar a expressão das coordenadas principais e padronizadas.

#### 2.4.4.1 ESPAÇO VETORIAL NÃO CENTRALIZADO

Analogamente à seção 2.3.4, a matriz que define o ajuste do subespaço no espaço não centralizado é dada por

$$Z^* = Y^{*'} Y^* \quad (2.4.24)$$

onde

$$\mathbf{Y}^* = (\mathbf{D}_c^{-1/2} \mathbf{C} \mathbf{D}_{\bar{c}}^{-1/2}) \quad (2.4.25)$$

As relações também são análogas para o vetor  $\mathbf{u}_a = \left( \sqrt{\frac{n_{1\bullet}}{n}}, \dots, \sqrt{\frac{n_{I\bullet}}{n}} \right)' = \mathbf{D}_{\bar{c}}^{-1/2} \mathbf{1}$  e,

portanto, as coordenadas principais dos perfis coluna podem, agora, ser expressas por

$$\mathbf{G} = \mathbf{D}_c^{-1/2} \mathbf{Y}^* \mathbf{U} \quad (2.4.26)$$

ou

$$\mathbf{G} = \mathbf{C} \mathbf{D}_{\bar{c}}^{-1/2} \mathbf{U}. \quad (2.4.27)$$

O subespaço que melhor se ajusta aos dados, em termos dos quadrados mínimos ponderados, é definido pelos  $k$  autovetores  $\mathbf{u}_k$ ,  $1 \leq k \leq K$ , correspondentes aos  $k$  maiores autovalores  $\eta_k$  de  $\mathbf{Z}$ . Para não sobrecarregar em notações, expressaremos as coordenadas dos subespaços ajustado e ótimo pela mesma expressão dada em (2.4.22) e (2.4.27), ficando portanto as dimensões das matrizes definidas pelo contexto.

Estudadas as relações entre as linhas e entre as colunas, dediquemo-nos ao estudo das relações entre linhas e colunas.

## 2.5 RELAÇÕES ENTRE LINHAS E COLUNAS

Vimos, até aqui, as coordenadas dos perfis linha e dos perfis coluna nos seus respectivos subespaços ótimos. Para isso, definimos: duas nuvens de pontos (uma em cada espaço), a métrica e encontramos os subespaços ótimos. Agora vamos ver as relações existentes entre estes dois subespaços, mostrando, assim, a correspondência entre as duas nuvens de pontos. Para mostrar a correspondência entre a nuvem  $\mathbf{R}$  em  $\mathbf{R}^J$  e a nuvem  $\mathbf{C}$  em  $\mathbf{R}^I$  mostramos, nesta seção, que são basicamente duas as relações entre elas: a relação entre o centróide e o vetor de massas e a relação entre as bases dos seus respectivos espaços vetoriais.

### 2.5.1 MASSAS E CENTRÓIDES

Primeiramente, vamos verificar que o centróide da nuvem dos perfis linha coincide com o vetor de massas dos perfis coluna e que o centróide da nuvem dos perfis coluna coincide com o vetor de massas dos perfis linha, isto é:

$$\bar{\mathbf{R}} = \mathbf{c} \text{ e } \bar{\mathbf{C}} = \mathbf{r}. \quad (2.5.1)$$

Vamos verificar então que  $\bar{\mathbf{R}} = \mathbf{c}$ , pois, como dado em (2.3.7) e (2.2.6) temos que

$$\bar{\mathbf{R}} = \mathbf{R}'\mathbf{r} \text{ e } \mathbf{c} = \mathbf{P}'\mathbf{1}_i$$

assim, substituindo a expressão matricial de  $\mathbf{R}$  dada em (2.3.3) em  $\bar{\mathbf{R}} = \mathbf{R}'\mathbf{r}$  temos que  $\bar{\mathbf{R}} = \mathbf{P}'\mathbf{D}_r^{-1}\mathbf{r}$ , mas como  $\mathbf{D}_r^{-1}\mathbf{r} = \mathbf{1}_i$ , temos  $\bar{\mathbf{R}} = \mathbf{P}'\mathbf{1}_i$ , e portanto  $\bar{\mathbf{R}} = \mathbf{c}$ . De forma análoga verifica-se que  $\bar{\mathbf{C}} = \mathbf{r}$ .

Conforme observação feita ao definir a matriz de correspondência em (2.2.4), ao expressarmos matematicamente as relações entre as nuvens, em termos da matriz da correspondência, temos as simplificações que nos facilitam a observação dessas relações. Assim, expressando as matrizes  $\mathbf{X}$  e  $\mathbf{Y}$  dadas em (2.3.22) e (2.4.19) em termos da matriz de correspondência, isto é, substituindo  $\mathbf{R}$  e  $\mathbf{C}$  pelas expressões matriciais dadas em (2.3.3) e (2.4.3) e usando as relações dadas em (2.5.1), temos

$$\mathbf{X} = \mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{r}\mathbf{c}') \mathbf{D}_c^{-1/2}, \quad (2.5.2)$$

$$\mathbf{Y} = \mathbf{D}_c^{-1/2}(\mathbf{P}' - \mathbf{c}\mathbf{r}') \mathbf{D}_r^{-1/2}. \quad (2.5.3)$$

Note que, agora, podemos observar diretamente que  $Y$  é exatamente a matriz transposta da matriz  $X$ , isto é,

$$Y = X' \quad (2.5.4)$$

e a matriz  $Z$  é, portanto, a transposta da matrizes  $S$ , pois de (2.5.4) temos que

$$S = X'X \text{ e } Z = X X' . \quad (2.5.5)$$

Com esta relação podemos, então, verificar a relação entre as bases dos espaços vetoriais. Temos ainda que a inércia total das linhas dada em (2.3.18) é igual à inércia total das colunas dada em (2.4.16), isto é,

$$In_{\bar{R}}(\mathbf{R}) = tr[X'X] = tr[XX'] = In_{\bar{C}}(\mathbf{C}) . \quad (2.5.6)$$

Uma outra relação interessante para se colocar aqui é que sendo  $S$  a matriz dos produtos internos entre as colunas de  $N$ , como expresso em (2.3.19), temos, na diagonal principal, as inércias das categorias das colunas. E a matriz  $Z$  tem, na diagonal principal, a inércia das categorias das linhas. Mas a expressão acima, a (2.5.6), nos mostra que a soma das inércias das linhas é igual à soma das inércias das colunas. Desta forma, as matrizes  $S$  e  $Z$  são, às vezes, chamadas de Matrizes de Inércias das colunas e das linhas, respectivamente.

Passamos a discutir a relação entre as coordenadas dos espaços vetoriais.

## 2.5.2 ESPAÇOS VETORIAIS

Vimos em (2.5.5) a relação entre as matrizes que definem os ajustes e os vetores direcionais que geram os dois subespaços; e em (2.3.31) e (2.4.20) as decomposições espectrais dessas matrizes. Reescrevendo-as temos

$$X'X V = V D_{\lambda} \text{ com } V'V = I \quad (2.5.7)$$

$$XX'U = U D_{\eta} \text{ com } U'U = I \quad (2.5.8)$$

onde  $V$  e  $U$  são as matrizes em cujas colunas estão os vetores direcionais que formam as bases dos espaços  $R^J$  e  $R^I$ , respectivamente. Fazendo operações de pré e pós-multiplicações nas expressões acima, encontramos a relação entre  $U$  e  $V$ . Esta relação pode ser obtida de duas maneiras.

Primeiramente, considerando a análise dos perfis linha, fazendo pré-multiplicações de  $X$  em (2.5.7) e comparando com (2.5.8) temos que  $D_{\eta} = D_{\lambda}^{(1)}$  e  $XV$  são os autovetores de  $XX'$

<sup>(1)</sup> ver Lebart, Morineau e Warwick(1984) pp 7.

associados aos autovalores  $\lambda_1, \dots, \lambda_J$ . Se os  $\lambda_k$  para  $k=1, \dots, K$  ( $\lambda_k \neq 0$ ) são todos distintos, então a direção dos autovalores de  $\mathbf{XX}'$  é definida de forma única, e portanto  $\mathbf{XV}$  é proporcional a  $\mathbf{U}$ , isto é, existe  $\mathbf{D}_a$  tal que

$$\mathbf{XV D}_a = \mathbf{U}. \quad (2.5.9)$$

Fazendo operações sobre (2.5.9) de pré-multiplicação por  $\mathbf{U}'$ , pré e pós-multiplicação por  $\mathbf{D}_a^{-1}$ , pré-multiplicação por  $\mathbf{V}$  e comparando com (2.5.8) temos

$$\mathbf{D}_a = \mathbf{D}_\lambda^{1/2}. \quad (2.5.10)$$

Portanto, de (2.5.9) temos a relação entre os dois subespaço vetoriais dados por

$$\mathbf{X} = \mathbf{U D}_\lambda^{-1/2} \mathbf{V}' \text{ com } \mathbf{U}'\mathbf{U}=\mathbf{V}'\mathbf{V}=\mathbf{I} \quad (2.5.11)$$

que é a *decomposição em valores singulares* da matriz  $\mathbf{X}$ .

Por outro lado este resultado também pode ser obtido considerando a análise dos perfis coluna, fazendo operações análogas e simétricas. A expressão (2.5.11) nos dá a relação entre os espaços dos perfis linha e dos perfis coluna. Essa relação nos mostra que só precisamos fazer uma análise, obtendo assim as bases dos dois espaços e, conseqüentemente, as suas coordenadas.

A expressão da decomposição em valores singulares de  $\mathbf{X}$  nos dá a relação entre as coordenadas dos pontos nas duas análises, as chamadas fórmulas de transição. A seguir veremos como obter as fórmulas de transição em coordenadas principais e em coordenadas padronizadas.

### 2.5.3 FÓRMULAS DE TRANSIÇÃO

Vamos verificar agora como as coordenadas desses dois subespaços estão relacionadas; pós-multiplicando (2.5.11) por  $\mathbf{V}$  e pré-multiplicando por  $\mathbf{D}_r^{-1/2}$  temos

$$\mathbf{D}_r^{-1/2} \mathbf{XV} = \mathbf{D}_r^{-1/2} \mathbf{U D}_\lambda^{-1/2}$$

que, substituindo pela expressão das coordenadas principais dos perfis linha em (2.3.33) e isolando  $\mathbf{U}$ , resulta em

$$\mathbf{U} = \mathbf{D}_r^{1/2} \mathbf{F D}_\lambda^{1/2}. \quad (2.5.12)$$

Da mesma forma, pré-multiplicando (2.5.11) por  $\mathbf{U}'$ , pós-multiplicando por  $\mathbf{D}_c^{-1/2}$ , substituindo pela expressão das coordenadas principais dos perfis coluna em (2.4.26) e isolando  $\mathbf{V}$ , temos

$$\mathbf{V} = \mathbf{D}_c^{1/2} \mathbf{G D}_\lambda^{1/2} \quad (2.5.13)$$

Substituindo as expressões de  $U$  e  $V$  obtidas em (2.5.12) e (2.5.13), respectivamente em (2.3.41) e (2.4.27), resulta

$$\mathbf{F} = \mathbf{R} \mathbf{G} \mathbf{D}_\lambda^{-1/2} \quad (2.5.14)$$

$$\mathbf{G} = \mathbf{C} \mathbf{F} \mathbf{D}_\lambda^{-1/2} \quad (2.5.15)$$

que são chamadas de *fórmulas de transição*. Elas nos dão as relações entre as coordenadas principais dos dois espaços.

Substituindo as coordenadas padronizadas em (2.3.41) e (2.4.27) em (2.5.14) e (2.5.15) respectivamente temos

$$\mathbf{F} = \mathbf{R} \mathbf{\Gamma} \quad (2.5.16)$$

$$\mathbf{G} = \mathbf{C} \mathbf{\Phi} \quad (2.5.17)$$

que são as fórmulas de transição entre espaços com coordenadas principais e padronizadas.

Obtidas as expressões das Coordenadas Principais e Padronizadas e também as fórmulas de transição que nos dão a relação entre as duas nuvens, vamos então ver como se faz a interpretação desses pontos no subespaço ótimo. Ou melhor, veremos as regras básicas para a interpretação dos gráficos da Análise de Correspondência.

## 2.6 INTERPRETAÇÃO

Ao aplicarmos o algoritmo da AC em tabelas de contingências, obtemos as coordenadas dos perfis linha e dos perfis coluna em coordenadas principais e em coordenadas padronizadas. Assim, de acordo com o interesse do estudo, aloca-se estas coordenadas no subespaço ótimo para interpretar graficamente as associações entre as linhas e as colunas desta tabela. Nesses gráficos são alocadas, conjuntamente, as coordenadas dos perfis linha e dos perfis coluna, que podem ser de três formas: a) os perfis linha em coordenadas principais e os perfis coluna também em coordenadas principais; b) os perfis linha em coordenadas principais e os perfis coluna em coordenadas padronizadas e c) os perfis coluna em coordenadas principais e os perfis linha em coordenadas padronizadas.

Para a obtenção dos eixos a serem analisados, observamos os percentuais acumulados das inércias principais e definimos quantos eixos (dimensões) são necessários para a representação dos pontos com uma precisão considerada satisfatória e que seja possível de ser visualizada, isto é, uma, duas ou três dimensões.

Tradicionalmente a interpretação era feita no plano; quando eram necessários mais eixos, faziam-se vários planos, por exemplo,  $dim1 \times dim2$ ,  $dim2 \times dim3$  e  $dim1 \times dim3$ . Mas hoje, com a evolução dos programas gráficos, podemos, quando necessário, analisar o gráfico tridimensional, que nos mostra com mais precisão as posições relativas dos pontos no espaço.

Para interpretar esses gráficos, primeiramente estudamos os pontos em cada eixo separadamente e depois estudamos os pontos no conjunto dos eixos. No estudo de cada eixo, estudamos a contribuição de cada ponto para a determinação desse eixo e a representabilidade de cada ponto neste eixo. No estudo do gráfico -o conjunto dos eixos- estudamos a proximidade dos pontos de uma mesma nuvem e as relações dadas pelas fórmulas de transição.

### 2.6.1 INTERPRETAÇÃO DOS EIXOS

Para a interpretação dos eixos, vamos verificar como a inércia dos pontos se decompõe em cada eixo. Isto é importante, pois são as componentes da inércia que formam os coeficientes que indicam o quanto cada ponto contribuiu para a determinação da direção dos eixos e o quanto cada

ponto está representado em cada eixo. Estes coeficientes foram denominados por Greenacre (1984) de contribuição absoluta e contribuição relativa de cada ponto.

Observe na figura 3 abaixo, por exemplo, um ponto no “simplex” transformado dos perfis linha (para  $J=3 < I$ ). Suponha que  $v_1$  e  $v_2$  sejam os eixos que indicam as direções de maior inércia da nuvem do ponto  $R_i$  e portanto  $f_{i1}$  e  $f_{i2}$  são as projeções desse ponto nos respectivos eixos, pois são elementos da matriz  $F$  dada em (2.3.33). Então, a inércia de cada uma destas projeções é dada por  $r_i(f_{i1})^2$  e  $r_i(f_{i2})^2$ , que são as componentes da inércia desse ponto nesses eixos. Elas são denotadas por  $[In_c(R_i)]_1$  e  $[In_c(R_i)]_2$ . Se este é o espaço ajustado, então, pelo Teorema de Pitágoras, a inércia deste ponto é a soma de suas componentes, isto é,

$$[In_c(R_i)]_1 + [In_c(R_i)]_2 = r_i(f_{i1})^2 + r_i(f_{i2})^2 = r_i[(f_{i1})^2 + (f_{i2})^2] = r_i d_i^2(R_i, c) = In_c(R_i).$$

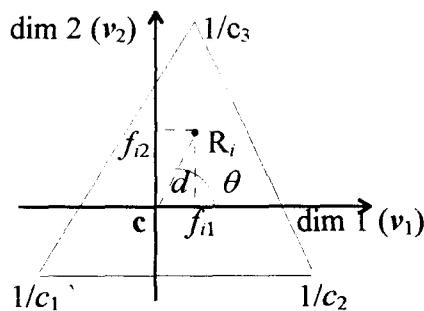


Figura 3: Decomposição da distância.

Desta forma, para um subespaço ajustado de dimensão  $K$ , a decomposição da inércia para o conjunto de pontos que representam as linhas da matriz de dados é mostrada na tabela 2 e a decomposição da inércia para o conjunto de pontos que representam as colunas é mostrada na tabela 3. Temos, nas linhas dessas tabelas, a inércia de cada ponto decomposta sobre cada eixo. Na última coluna de cada tabela, temos as somas, com respeito aos eixos, das componentes da inércia de cada ponto. Estas somas, em cada linha, nos dão a inércia do ponto. Na última linha, temos a soma das componentes da inércia de todos os pontos, num mesmo eixo. Aqui estas somas nos dão a inércia principal de cada eixo. Lembramos que estas inércias principais são também dadas pelos autovalores da matriz que definem o ajuste das nuvens. A soma dos elementos da última linha ou os da última coluna nos dá a inércia total das nuvens de pontos.

TABELA 2: Decomposição da Inércia dos perfis linha sobre os eixos principais.

|              |          | Eixos          |     |                | Total              |
|--------------|----------|----------------|-----|----------------|--------------------|
|              |          | $v_1$          | ... | $v_K$          |                    |
| Perfis linha | $R_1$    | $r_1 f_{11}^2$ | ... | $r_1 f_{1K}^2$ | $In_c(R_1)$        |
|              | $\vdots$ | $\vdots$       |     | $\vdots$       | $\vdots$           |
|              | $R_I$    | $r_I f_{I1}^2$ | ... | $r_I f_{IK}^2$ | $In_c(R_I)$        |
| Total        |          | $\lambda_1$    | ... | $\lambda_K$    | $In_c(\mathbf{R})$ |

$$\text{onde } \lambda_k = \sum_{i=1}^I r_i f_{ik}^2, k = 1, \dots, K \text{ e } In_c(R_i) = r_i \sum_{k=1}^K f_{ik}^2, i = 1, \dots, I$$

TABELA 3: Decomposição da Inércia dos perfis coluna sobre os eixos principais.

|               |          | Eixos          |     |                | Total              |
|---------------|----------|----------------|-----|----------------|--------------------|
|               |          | $u_1$          | ... | $u_K$          |                    |
| Perfis coluna | $C_1$    | $c_1 g_{11}^2$ | ... | $c_1 g_{1K}^2$ | $In_r(C_1)$        |
|               | $\vdots$ | $\vdots$       |     | $\vdots$       | $\vdots$           |
|               | $C_J$    | $c_J g_{J1}^2$ | ... | $c_J g_{JK}^2$ | $In_r(C_J)$        |
| Total         |          | $\lambda_1$    | ... | $\lambda_K$    | $In_r(\mathbf{C})$ |

$$\text{onde } \lambda_k = \sum_{j=1}^J c_j g_{jk}^2, k = 1, \dots, K \text{ e } In_r(C_j) = c_j \sum_{k=1}^K g_{jk}^2, j = 1, \dots, J.$$

Para interpretar as relações entre linhas e entre colunas de uma tabela de contingência sobre os subespaços gerados pelos eixos obtidos pela análise de correspondência é necessário analisar dois conjuntos de coeficientes para cada um dos perfis analisados em cada eixo:

1) A *Contribuição Absoluta*, que exprime o quanto um perfil contribui para a inércia principal do eixo em estudo, é o quociente entre a componente da inércia do perfil no eixo e a inércia principal do eixo, dado por

$$ca[R_i]_k = [In_c(R_i)]_k / \sum_k [In_c(R_i)]_k = r_i f_{ik}^2 / \lambda_k \quad (2.6.1)$$

$$= r_i \phi_{ik}^2$$

ou

$$ca[C_j]_k = [In_r(C_j)]_k / \sum_k [In_r(C_j)]_k = c_j g_{jk}^2 / \lambda_k \quad (2.6.2)$$

$$= c_j \gamma_{ik}^2.$$

Observe que é a partir da contribuição absoluta (quociente entre a componente da inércia dos pontos em cada eixo e a sua inércia principal) que definimos as coordenadas padronizadas, como expresso em (2.3.34), pois quando dividimos, por exemplo,  $[In_c(R_i)]_k / \lambda_k$  obtemos  $r_i f_{ik}^2 / \lambda_k = r_i (f_{ik} / \lambda^{1/2})^2$  e portanto  $f_{ik} / \lambda^{1/2} = \phi_{ik}$  é o que chamamos de coordenada padronizada.

2) A **Contribuição Relativa**, que exprime o quanto o perfil está representado neste eixo, é o quociente entre a componente da inércia do perfil no eixo e a inércia do perfil. Dai o nome contribuição relativa, pois ela é a contribuição da inércia do perfil ao eixo, em relação à inércia do perfil, que mede a qualidade de representação do perfis pelo eixo, dado por

$$\begin{aligned} cr[R_i]_k &= [In_c(R_i)]_k / In_c(R_i) = f_{ik}^2 / \sum_k f_{ik}^2 & (2.6.3) \\ &= \phi_{ik}^2 / \sum_k \phi_{ik}^2 \end{aligned}$$

ou

$$\begin{aligned} cr[C_j]_k &= [In_r(C_j)]_k / In_r(C_j) = g_{jk}^2 / \sum_k g_{jk}^2 & (2.6.4) \\ &= \gamma_{jk}^2 / \sum_k \gamma_{jk}^2 \end{aligned}$$

Observe que a contribuição relativa é equivalente a tomar o quadrado do coseno do ângulo entre o vetor direcional do ponto e o eixo, pois, ao dividirmos a componente da inércia do perfil pela sua inércia, temos o cancelamento da massa, resultando no quadrado do quociente entre a coordenada desse ponto no eixo e a distância do ponto ao centro, isto é,

$$[In_c(R_i)]_k / In_c(R_i) = r_i f_{ik}^2 / r_i \sum_k f_{ik}^2 = f_{ik}^2 / \sum_k f_{ik}^2 = f_{ik}^2 / d^2(R_i, \mathbf{c}) = (f_{ik} / d(R_i, \mathbf{c}))^2 = (\cos(\theta))^2$$

onde  $\theta$  é o ângulo entre o vetor direcional do ponto  $R_i$  e o eixo  $v_k$ , como podemos observar na figura 3, para  $K=2$ .

A interpretação dos eixos é fundamental, uma vez que a visão direta do gráfico pode induzir a erros. Para evitar esses erros, será necessário levar em conta as contribuições absolutas e relativas dos pontos. A seguir apresentamos um roteiro para a interpretação em cada eixo:

- a) Buscam-se os pontos que apresentam maior contribuição absoluta (*ca*) no eixo considerado.
- b) Destacam-se, entre estes pontos, aqueles que estão em oposição com relação à origem.
- c) Estuda-se a qualidade de representação destes pontos. Se um ponto tem uma contribuição relativa (*cr*) pequena, é de se supor que tenha um papel importante sobre outro eixo; assim, para seu estudo, será necessário considerar o conjunto dos eixos; buscam-se, então, aqueles pontos que estejam bem representados, isto é, que têm uma contribuição relativa (*cr*) alta. Estes pontos são ilustrativos da significação do eixo.

Desde que o eixo é uma nova dimensão que condensa e resume critérios inter-relacionados, pode-se, às vezes, dar significado a eles. Isso é relativamente fácil quando há elementos muito bem representados sobre o eixo e opostos entre si. Em geral se interpretam os eixos por oposição de adjetivos.

### 2.6.2 INTERPRETAÇÃO DOS PONTOS NO CONJUNTO DOS EIXOS

Interpretados cada eixo, vamos ver como se interpretam os pontos no gráfico propriamente dito que é formado pelo conjunto dos eixos que foram considerados. Esses eixos são os mais representativos da análise, isto é, os de maiores inércias principais de cada nuvem. Eles são superpostos para a representação conjunta das duas nuvens de pontos, que podem estar ambas em coordenadas principais ou uma em coordenadas principais e a outra em coordenadas padronizadas. Para qualquer dos três casos, primeiramente estuda-se a qualidade de representação dos pontos no subespaço considerado.

A *qualidade de representação* ( $qr$ ) é o coeficiente que nos dá o quanto cada ponto está ou não bem representado nesse subespaço (subespaço ótimo). Tal coeficiente é dado pela soma das contribuições relativas do ponto em cada eixo que, por sua vez, é dada por

$$qr[R_i] = \sum_{k=1}^{K^*} cr[R_i]_k \quad (2.6.5)$$

ou

$$qr[C_j] = \sum_{k=1}^{K^*} cr[C_j]_k \quad (2.6.6)$$

onde  $K^*$  é a dimensão do subespaço ótimo. Lembrando que, se o subespaço ótimo coincide com o subespaço ajustado, temos todos os pontos com  $qr = 1$ , ou seja, todos os pontos estarão 100% representados no subespaço ótimo.

Depois de estudar a qualidade de representação dos pontos, estuda-se a massa desses pontos para compreender o seu deslocamento em relação à origem. Aí então partimos para o estudo da proximidade entre eles.

Não é possível interpretar a proximidade entre dois pontos de nuvens diferentes, uma vez que não foi estabelecida nenhuma medida de distância entre estes pontos. A análise de pontos entre nuvens diferentes se dá através da interpretação das fórmulas de transição.

Quando representamos os pontos em coordenadas principais, temos alguns resultados com relação às posições entre eles, como apresentado em Villalobos Aguayo (1993) e que serão colocados a seguir.

1) Se duas linhas têm uma estrutura semelhante, isto é, se seus perfis são similares, então ambas as linhas relacionam-se de modo similar com as colunas. Portanto, no gráfico, seus pontos estarão localizados próximos um do outro. Mas cabe advertir que a inversa não é sempre verdadeira; isto dependerá da qualidade de representação dos pontos. Para as colunas, o resultado é análogo.

2) Quando dois pontos linha estão muito distanciados, então, suas linhas estão relacionadas de modos diferentes com as colunas. Quando dois pontos linha estão em direções opostas do centro, eles desviam-se de formas opostas do perfil médio. Para as colunas, o resultado é análogo.

Quando representamos as linhas e colunas simultaneamente, sobre o mesmo gráfico, em termos das coordenadas principais, devemos considerar as fórmulas de transição dadas em (2.5.14) e (2.5.15). Reescrevendo a primeira delas, temos

$$\mathbf{F} = \mathbf{R}\mathbf{G}\mathbf{D}_\lambda^{-1/2}$$

$$F_i = \left( \frac{1}{\lambda_1^{1/2}} \sum_{j=1}^J R_{ij} g_{j1}, \frac{1}{\lambda_2^{1/2}} \sum_{j=1}^J R_{ij} g_{j2}, \dots, \frac{1}{\lambda_K^{1/2}} \sum_{j=1}^J R_{ij} g_{jK} \right)' = \mathbf{D}_\lambda^{-1/2} \left[ \sum_{j=1}^J R_{ij} \mathbf{G}_j \right] \quad (2.6.7)$$

A partir desta fórmula, podemos interpretar as posições das linhas com relação às posições das colunas, da seguinte forma:

1) Uma linha  $i$  estará mais atraída na direção de uma coluna  $j$ , quanto maior for o  $j$ -ésimo elemento do perfil linha  $i$ . Isto ocorre porque as coordenadas do ponto que representa a linha  $i$ , o vetor  $F_i$ , é resultante de uma soma ponderada das coordenadas dos pontos que representam as colunas, onde os elementos ponderadores das  $j$  colunas são os elementos do perfil linha  $i$ , e portanto se  $j$ -ésimo elemento do perfil linha  $i$  é maior, o vetor  $F_i$  apresenta uma proporção maior do vetor  $\mathbf{D}_\lambda^{-1/2} \mathbf{G}_j$ , como podemos observar em (2.6.7).

2) A proximidade de uma linha  $i$  a uma coluna  $j$  só pode ser interpretada como uma associação entre elas, se ambos os pontos estiverem longe da origem, na periferia da nuvem. Caso contrário, duas situações podem acontecer sem que exista associação: a) Quando uma linha tem um perfil próximo ao perfil médio, se encontra próxima à origem; de modo que a linha não

desempenha um papel importante para nenhuma das colunas, já que se distribuem sem uma tendência particular, tal como o faz o total do grupo observado. b) A proximidade entre um ponto que representa uma linha  $i$  e um ponto que representa uma coluna  $j$  poderia dever-se não ao fato de que a coluna  $j$  tenha uma importância sobre a linha  $i$ . Mas sim ao fato de que mais de uma coluna que tem importância sobre a linha  $i$  se encontra muito longe da coluna  $j$  e em extremos opostos.

3) A posição de um ponto linha deve ser interpretada sempre com relação às posições de todos os pontos coluna.

A interpretação das posições das colunas com relação às linhas é feita com base na fórmula de transição (2.5.15) e os resultados serão análogos aos anteriores.

A relação entre os dois tipos de coordenadas é expressa pelas fórmulas de transição dadas em (2.5.16) e (2.5.17), que reescrevemos a seguir.

$$\mathbf{F} = \mathbf{R} \Gamma \text{ e } \mathbf{G} = \mathbf{C} \Phi$$

A primeira fórmula nos diz que a projeção do ponto que representa uma linha  $i$  sobre o eixo  $k$  é exatamente o baricentro das projeções das  $j$  colunas,  $\Gamma_j$ , sobre o mesmo eixo. Cada coluna está afetada por um peso que é a sua importância relativa na linha  $i$ . A segunda fórmula nos dá, de forma análoga e simétrica, o mesmo resultado para as colunas.

Quando representarmos, nos gráficos, uma nuvem de pontos em coordenadas principais e a outra em coordenadas padronizadas, nossa interpretação pode ser grandemente facilitada. Por exemplo, num gráfico onde as linhas estão representadas pelas coordenadas principais e as colunas estão representadas pelas coordenadas padronizadas, a leitura fica bem mais direta, pois cada perfil linha é exatamente o baricentro dos pontos coluna, onde os pesos são os elementos do respectivo perfil linha. Como podemos observar ao reescrevermos um vetor da matriz dada em (2.5.16).

$$F_i = \sum_{j=1}^J R_{ij} \Gamma_j$$

Neste caso veremos que, quanto maior for o  $j$ -ésimo elemento do perfil linha  $i$ ,  $R_{ij}$ , haverá uma maior proximidade do ponto que representa a linha  $i$  ao ponto que representa a coluna  $j$ . Isto acontece porque o vetor resultante  $F_i$  contém uma porção maior do vetor  $\Gamma_j$ . Também podemos

concluir que, devido à ponderação do espaço feita no início, quanto menor a massa da coluna  $j$ , mais longe se encontrará o ponto  $\Gamma_j$  do centróide.

No caso em que as linhas estão representadas em coordenadas padronizadas e as colunas em coordenadas principais, o resultado é análogo.

Desta forma, interpretados os pontos que contribuíram para a determinação dos eixos, podemos então interpretar os pontos suplementares, se for o caso.

### 2.6.3 PONTOS SUPLEMENTARES

Uma vez que os eixos principais das nuvens de pontos foram estabelecidos, é possível representar pontos adicionais no espaço definido pelos perfis. O conceito de ponto suplementar é muito importante na análise de correspondência. Esses pontos adicionais podem ser vistos como pontos com massa igual a zero, uma vez que eles não participam da definição da distância, nem na determinação dos eixos principais.

As coordenadas de uma linha (ou coluna) suplementar  $s$ , de valores  $n_{sj}$  ( $n_{is}$ )  $i = 1, \dots, I$ , e  $j = 1, \dots, J$ , são obtidas por meio das fórmulas de transição:

$$F_s = R_s' G D_\lambda^{-1} \quad (2.6.8)$$

ou

$$G_s = C_s' F D_\lambda^{-1} \quad (2.6.9)$$

onde  $R_s$  e  $G_s$  são, respectivamente, o perfil linha e o perfil coluna de uma linha e de uma coluna suplementar.

Uma vantagem de se mostrar tais pontos é que eles podem melhorar a interpretação dos eixos principais e dos padrões ou estruturas observadas nos perfis. O uso dos pontos suplementares é muito útil e às vezes, em alguns casos, necessário. Geralmente são utilizados para: a) representar os centróides de algumas classes, ou, ao contrário, representar os elementos da classe, quando os eixos são determinados pelos centróides; b) representar a caracterização de distintos grupos de indivíduos; c) representar um elemento que perturba a análise anterior, por ser muito diferente, impedindo a interpretação clara da análise.

Finalizando, assim, o estudo da Análise de Correspondência Simples, no próximo capítulo discutiremos a Análise de Correspondência em Grupos.

## 3. ANÁLISE DE CORRESPONDÊNCIA EM GRUPOS

### 3.1 INTRODUÇÃO

Com muita frequência, encontramos dados onde se pode considerar uma estrutura de grupo nas categorias das variáveis e o interesse do estudo está voltado para as diferenças entre os grupos. Essas diferenças são observáveis através da investigação das associações existentes entre as categorias dentro de cada grupo e/ou entre as categorias de diferentes grupos. Podemos citar, como exemplo de onde se tem este interesse, o estudo feito por Greenacre e Vrba (1984) e, também, estudos envolvendo dados censitários, pelo interesse em diferenças regionais e temporais.

Existem diferentes maneiras de investigar associações como as acima colocadas. Uma opção é aplicarmos a análise de correspondência em cada um dos grupos separadamente e então compará-los. Deste modo, faz-se uma Análise de Correspondência Simples em cada grupo. Ocorre que, neste caso, os gráficos onde os pontos são representados não têm escalas necessariamente iguais, o que torna imprecisa esta comparação. O problema da diferença de escalas pode ser contornado ao se justapor as tabelas de contingência nas quais foi feita a análise de correspondência simples, efetuando assim a Análise de Correspondência Conjunta. Deste modo, todos os grupos terão os pontos dispostos num mesmo gráfico, e portanto em uma mesma escala. Mas, neste caso, surge um novo problema: os eixos direcionais nos mostram a direção de maior inércia da nuvem como um todo, não realçando as diferenças entre os grupos, podendo até mascará-las. Assim sendo, consideramos que a melhor opção para comparar os resultados entre os

grupos é tomar um ponto como representante de cada grupo e maximizar as suas inércias; deste modo estamos maximizando as diferenças entre os grupos. E, assim, estamos efetuando a Análise de Correspondência em Grupos.

Greenacre (1984), ao tomar um ponto como representante do grupo, escolhe o centróide. Neste trabalho não entramos na questão da escolha do representante de cada grupo. Em estudos de Análise Discriminante procuram encontrar um ponto ideal para esta representação, como faz Takane (1989). Aqui, estamos seguindo a sugestão de Greenacre, tomamos o centróide como sendo o representante de seu grupo. A partir da maximização das inércias dos centróides, maximizamos as diferenças entre os grupos e, neste caso, os eixos direcionais nos mostram a direção de maior inércia dos grupos. Com essa estrutura, e colocando os pontos de cada nuvem como pontos suplementares, podemos investigar as associações entre categorias dos diferentes grupos ou, ainda, comparar os resultados obtidos dentro de cada grupo.

Para estudar as associações entre categorias dos diferentes grupos e comparar os resultados obtidos dentro destes grupos, nesse capítulo desenvolvemos a AC para dados de uma tabela de contingência que apresente estrutura de grupo nas categorias das duas variáveis categóricas e mostramos alguns resultados que relacionam as análises das quatro matrizes associadas a esta tabela. Em cada uma delas, o desenvolvimento do algoritmo da AC é feito como no capítulo anterior. Porém, como cada linha ou coluna agora representa um grupo, esta análise fica equivalente à Análise Discriminante. Assim podemos dizer que a Análise de Correspondência em Grupos é uma dupla Análise Discriminante para dados categóricos, pois é feita uma análise nas linhas e outra nas colunas da tabela.

Na seção 3.2, mostramos como os dados de uma tabela de contingência com estrutura de grupo são apresentados. A esta tabela, é associada a matriz  $\mathbf{N}$ , que é a mesma matriz de dados onde agora é realçada a estrutura de grupo. A partir da matriz  $\mathbf{N}$ , obtemos três outras matrizes nas quais é feito o agrupamento das categorias em cada grupo das linhas e/ou colunas. Na seção 3.3, estudamos as matrizes que nos dão as associações existentes entre as linhas e os agrupamentos das colunas, entre as colunas e os agrupamentos das linhas, e entre os agrupamentos das linhas e das colunas. Neste estudo, aplicamos o algoritmo da AC nas tabelas acima e apresentamos os resultados que as relacionam.

**3.2 OS DADOS**

Na Análise de Correspondência em Grupos, consideramos a Análise de Correspondência como a técnica que exhibe os perfis das  $I$  categorias da variável das linhas e das  $J$  categorias da variável das colunas da tabela 1, com respeito a um espaço ótimo, mas com a atenção voltada para as partições das categorias ou dos indivíduos de cada variável categórica.

Neste capítulo estamos considerando apenas uma partição em cada uma das variáveis. Portanto, assumiremos que as categorias  $i = 1, \dots, I$ , estão agrupadas em  $H$  classes disjuntas e

exaustivas, tais que  $I = \sum_{h=1}^H I_h$ , onde  $I_h$  é o número de categorias no  $h$ -ésimo grupo das linhas, o

mesmo ocorrendo com as categorias  $j = 1, \dots, J$ , que são agrupadas em  $L$  classes disjuntas e

exaustivas, tais que  $J = \sum_{\ell=1}^L J_\ell$  onde  $J_\ell$  é o número de categorias no  $\ell$ -ésimo grupo das colunas.

Assim, os dados em estudo neste capítulo são apresentados em uma tabela de contingência na qual a estrutura de grupos é como a da tabela 4.

Tabela 4: Modelo de tabela de contingência com estrutura de grupos nas duas variáveis

|                 |         | Variável Categórica B       |                      |                               |                                  |                           |                             |                      |                               |                                  |                                       |                                 |
|-----------------|---------|-----------------------------|----------------------|-------------------------------|----------------------------------|---------------------------|-----------------------------|----------------------|-------------------------------|----------------------------------|---------------------------------------|---------------------------------|
|                 |         | Grupo 1                     |                      |                               |                                  |                           | ...                         | Grupo L              |                               |                                  |                                       | Total                           |
|                 |         | cat 1                       | ...                  | cat $J_1$                     | total                            |                           | cat 1                       | ...                  | cat $J_L$                     | total                            |                                       |                                 |
| Variável        | Grupo 1 | cat. 1                      | $n_{11}^{11}$        | ...                           | $n_{1J_1}^{11}$                  | $n_{1\bullet}^{11}$       |                             | $n_{11}^{1L}$        | ...                           | $n_{1J_L}^{1L}$                  | $n_{1\bullet}^{1L}$                   | $n_{i\bullet}^{1\bullet}$       |
|                 |         | ...                         | ...                  | ...                           | ...                              | ...                       | ...                         | ...                  | ...                           | ...                              | ...                                   | ...                             |
|                 |         | cat $I_1$                   | $n_{I_11}^{11}$      | ...                           | $n_{I_1J_1}^{11}$                | $n_{I_1\bullet}^{11}$     |                             | $n_{I_11}^{1L}$      | ...                           | $n_{I_1J_L}^{1L}$                | $n_{I_1\bullet}^{1L}$                 | $n_{I_1\bullet}^{1\bullet}$     |
|                 |         | total                       | $n_{\bullet 1}^{11}$ | ...                           | $n_{\bullet J_1}^{11}$           | $n_{\bullet\bullet}^{11}$ |                             | $n_{\bullet 1}^{1L}$ | ...                           | $n_{\bullet J_L}^{1L}$           | $n_{\bullet\bullet}^{1L}$             | $n_{\bullet\bullet}^{1\bullet}$ |
| Cate-<br>górica | Grupo H | cat 1                       | $n_{11}^{H1}$        | ...                           | $n_{1J_1}^{H1}$                  | $n_{1\bullet}^{H1}$       |                             | $n_{11}^{HL}$        | ...                           | $n_{1J_L}^{HL}$                  | $n_{1\bullet}^{HL}$                   | $n_{i\bullet}^{H\bullet}$       |
|                 |         | ...                         | ...                  | ...                           | ...                              | ...                       | ...                         | ...                  | ...                           | ...                              | ...                                   |                                 |
|                 |         | cat $I_H$                   | $n_{I_H 1}^{H1}$     | ...                           | $n_{I_H J_1}^{H1}$               | $n_{I_H \bullet}^{H1}$    |                             | $n_{I_H 1}^{HL}$     | ...                           | $n_{I_H J_L}^{HL}$               | $n_{I_H \bullet}^{HL}$                | $n_{I_H \bullet}^{H\bullet}$    |
|                 |         | total                       | $n_{\bullet 1}^{H1}$ | ...                           | $n_{\bullet J_1}^{H1}$           | $n_{\bullet\bullet}^{H1}$ |                             | $n_{\bullet 1}^{HL}$ | ...                           | $n_{\bullet J_L}^{HL}$           | $n_{\bullet\bullet}^{HL}$             | $n_{\bullet\bullet}^{H\bullet}$ |
| Total           |         | $n_{\bullet 1}^{\bullet 1}$ | ...                  | $n_{\bullet J_1}^{\bullet 1}$ | $n_{\bullet\bullet}^{\bullet 1}$ | ...                       | $n_{\bullet 1}^{\bullet L}$ | ...                  | $n_{\bullet J_L}^{\bullet L}$ | $n_{\bullet\bullet}^{\bullet L}$ | $n_{\bullet\bullet}^{\bullet\bullet}$ |                                 |

onde:

$n_{ij}^{h\ell}$  é o número de indivíduos na  $i$ -ésima categoria do  $h$ -ésimo grupo da variável categórica A e na  $j$ -categoria do  $\ell$ -ésimo grupo da variável categórica B.

$n_{\bullet j}^{h\ell} = \sum_{i=1}^{I_h} n_{ij}^{h\ell}$  é o número total de indivíduos do  $h$ -ésimo grupo da variável categórica A, na  $j$ -ésima categoria do  $\ell$ -ésimo grupo da variável categórica B.

$n_{i\bullet}^{h\ell} = \sum_{j=1}^J n_{ij}^{h\ell}$  é o número total de indivíduos do  $\ell$ -ésimo grupo da variável categórica B, na  $i$ -ésima categoria do  $h$ -ésimo grupo da variável categórica A.

$n_{\bullet j}^{h\ell} = \sum_{h=1}^H n_{\bullet j}^{h\ell}$  é o número total de indivíduos da variável categórica A, na  $j$ -ésima categoria do  $\ell$ -ésimo grupo da variável categórica B.

$n_{i\bullet}^{h\ell} = \sum_{\ell=1}^L n_{i\bullet}^{h\ell}$  é o número total de indivíduos da variável categórica B, na  $i$ -ésima categoria do  $h$ -ésimo grupo da variável categórica A.

$n_{\bullet\bullet}^{h\ell} = \sum_{i=1}^{I_h} n_{i\bullet}^{h\ell} = \sum_{j=1}^J n_{\bullet j}^{h\ell}$  é o número total de indivíduos do  $h$ -ésimo grupo da variável categórica A, no  $\ell$ -ésimo grupo da variável categórica B.

$n_{\bullet\bullet}^{h\ell} = \sum_{\ell=1}^L n_{\bullet\bullet}^{h\ell}$  é o número total de indivíduos do  $h$ -ésimo grupo da variável categórica A, na variável categórica B.

$n_{\bullet\bullet}^{h\ell} = \sum_{h=1}^H n_{\bullet\bullet}^{h\ell}$  é o número total de indivíduos do  $\ell$ -ésimo grupo da variável categórica B, na variável categórica A.

$n_{\bullet\bullet} = \sum_{\ell=1}^L n_{\bullet\bullet}^{h\ell} = \sum_{h=1}^H n_{\bullet\bullet}^{h\ell} = n$  é o número total de indivíduos da tabela.

Associada à tabela 4, temos a matriz  $N$ , dada em (3.2.1), onde  $n_{ij}^{h\ell}$  é a interseção da  $i$ -ésima linha do  $h$ -ésimo grupo das linhas com a  $j$ -ésima coluna do  $\ell$ -ésimo grupo das colunas.

Da matriz de dados  $N$ , podemos obter três matrizes de onde fazemos as análises de acordo com interesse do problema. Sejam  $N_1$  -a matriz das freqüências das  $I$  linhas nos  $L$  grupos das colunas;  $N_2$  -a matriz das freqüências dos  $H$  grupos das linhas nas  $J$  colunas e  $N_3$  -a matriz das freqüências dos  $H$  grupos das linhas nos  $L$  grupos das colunas, dadas em (3.2.2), (3.2.3) e (3.2.4).

As frequências dos grupos das linhas e das colunas são obtidas quando se faz o *agrupamento* das linhas ou colunas destes grupos, em uma única linha ou coluna. Aqui, este agrupamento é feito somando-se as linhas ou colunas de cada grupo.

$$\mathbf{N} = \begin{bmatrix} n_{11}^{11} & \cdots & n_{1j_1}^{11} & \cdots & n_{11}^{1L} & \cdots & n_{1j_L}^{1L} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ n_{1_1}^{11} & \cdots & n_{1_1 j_1}^{11} & \cdots & n_{1_1}^{1L} & \cdots & n_{1_1 j_L}^{1L} \\ \vdots & & \vdots & & \vdots & & \vdots \\ n_{11}^{H1} & \cdots & n_{1j_1}^{H1} & \cdots & n_{11}^{HL} & \cdots & n_{1j_L}^{HL} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ n_{1_H}^{H1} & \cdots & n_{1_H j_1}^{H1} & \cdots & n_{1_H}^{HL} & \cdots & n_{1_H j_L}^{HL} \end{bmatrix} \quad (3.2.1)$$

$$\mathbf{N}_1 = \begin{bmatrix} n_{1\bullet}^{11} & \cdots & n_{1\bullet}^{1L} \\ \vdots & & \vdots \\ n_{1_1\bullet}^{11} & \cdots & n_{1_1\bullet}^{1L} \\ \vdots & & \vdots \\ n_{1\bullet}^{H1} & \cdots & n_{1\bullet}^{HL} \\ \vdots & & \vdots \\ n_{1_H\bullet}^{H1} & \cdots & n_{1_H\bullet}^{HL} \end{bmatrix} \quad (3.2.2)$$

$$\mathbf{N}_2 = \begin{bmatrix} n_{\bullet 1}^{11} & \cdots & n_{\bullet j_1}^{11} & \cdots & n_{\bullet 1}^{1L} & \cdots & n_{\bullet j_L}^{1L} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ n_{\bullet 1}^{H1} & \cdots & n_{\bullet j_1}^{H1} & \cdots & n_{\bullet 1}^{HL} & \cdots & n_{\bullet j_L}^{HL} \end{bmatrix} \quad (3.2.3)$$

$$\mathbf{N}_3 = \begin{bmatrix} n_{\bullet\bullet}^{11} & \cdots & n_{\bullet\bullet}^{1L} \\ \vdots & & \vdots \\ n_{\bullet\bullet}^{H1} & \cdots & n_{\bullet\bullet}^{HL} \end{bmatrix} \quad (3.2.4)$$

Algebricamente, essas matrizes podem ser obtidas ao definir uma matriz indicadora  $L \times J$  representada por  $\mathbf{Z}_0$ , cuja  $j$ -ésima coluna é um vetor de zeros, exceto por um "1" (um) na linha apropriada para indicar o grupo ao qual a coluna pertence e uma outra matriz indicadora  $I \times H$  representada por  $\tilde{\mathbf{Z}}_0$ , cujas linhas classificam, de forma similar, as linhas de  $\mathbf{N}$ . Assim a matriz  $\mathbf{N}_1$

é dada por  $N_1 = NZ_0'$ ,  $N_2$  é dada por  $N_2 = \tilde{Z}_0'N$  e  $N_3 = \tilde{Z}_0'N Z_0'$ . Todas essas matrizes estão representadas esquematicamente na figura 4 abaixo:

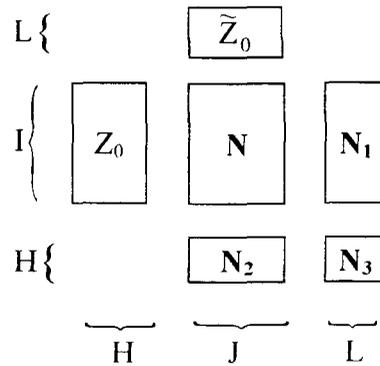


Figura 4 :Representação esquemática das matrizes  $N$ ,  $N_1$ ,  $N_2$  e  $N_3$

Apresentados os dados, passemos a estudar as associações entre as categorias e entre as classes das variáveis em questão.

### 3.3 RELAÇÕES ENTRE AS MATRIZES $N$ , $N_1$ , $N_2$ E $N_3$

Nesta seção, apresentamos as relações existentes entre as análises de correspondência das matrizes  $N$ ,  $N_1$ ,  $N_2$  e  $N_3$ . Salientando que a maioria dessas relações foram enunciadas, e algumas demonstradas, em Greenacre (1984) na seção 7.1, aqui elas serão desenvolvidas algebricamente, dentro de um contexto mais detalhado. Para isso, colocamos os resultados das análises dessas quatro matrizes e mostramos as relações entre elas.

#### 3.3.1 AS NUVENS DE PONTOS E SUAS RELAÇÕES

Quando a AC foi desenvolvida, no capítulo 2, apresentamos os perfis linha da matriz  $N$  na expressão (2.3.1). Tomando agora tal expressão e considerando a estrutura de grupo, ela é reescrita da seguinte maneira:

$$(R_i^h) = \left( \frac{n_{i1}^{h1}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{iJ_1}^{h1}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{i1}^{hL}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{iJ_L}^{hL}}{n_{i\bullet}^{h\bullet}} \right), i = 1, \dots, J_h \text{ e } h = 1, \dots, H. \quad (3.3.1)$$

Assim,  $(R_i^h)$  representa a  $i$ -ésima linha do  $h$ -ésimo grupo de  $N$  como um ponto no espaço  $\mathbf{R}^J$ .

Considerar estrutura de grupo na matriz  $N$  não muda a necessidade de se atribuir massa aos pontos, para que a informação sobre a proporção de indivíduos em cada linha seja conhecida. Deste modo, para a matriz  $N$ , temos que as massas desses pontos são:

$$(r_i^h) = \frac{n_{i\bullet}^{h\bullet}}{n}, i = 1, \dots, J_h \text{ e } h = 1, \dots, H \quad (3.3.2)$$

No caso das matrizes obtidas a partir do agrupamento das linhas e/ou colunas de  $N$ , os perfis linha e suas massas são escritos de forma similar.

Assim, os perfis linha de  $N_1$  são dados por:

$$(R_i^h)_1 = \left( \frac{n_{i\bullet}^{h1}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{i\bullet}^{hL}}{n_{i\bullet}^{h\bullet}} \right), i = 1, \dots, J_h \text{ e } h = 1, \dots, H \quad (3.3.3)$$

onde  $(R_i^h)_1$  representa o perfil linha da  $i$ -ésima linha do  $h$ -ésimo grupo de  $N_1$ , como um ponto no espaço  $\mathbf{R}^L$ . As massas dos perfis linha de  $N_1$  são dadas por

$$(r_i^h)_1 = \frac{n_{i\bullet}^{h\bullet}}{n}, i = 1, \dots, I_h \text{ e } h = 1, \dots, H. \quad (3.3.4)$$

Da mesma forma, os pontos do espaço  $\mathbf{R}^J$ , que representam os perfis linha de  $N_2$ , são dados por

$$(R_{\bullet}^h)_2 = \left( \frac{n_{\bullet 1}^{h1}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet J_1}^{h1}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet 1}^{hL}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet J_L}^{hL}}{n_{\bullet\bullet}^{h\bullet}} \right), h = 1, \dots, H \quad (3.3.5)$$

sendo que  $(R_{\bullet}^h)_2$  representa o perfil da  $h$ -ésima linha de  $N_2$ , enquanto que as massas desses perfis são:

$$(r_{\bullet}^h)_2 = \frac{n_{\bullet\bullet}^{h\bullet}}{n}, h = 1, \dots, H. \quad (3.3.6)$$

Finalmente, escreveremos a expressão que nos dá os perfis linha de  $N_3$ , que são pontos do espaço  $\mathbf{R}^L$ :

$$(R_{\bullet}^h)_3 = \left( \frac{n_{\bullet\bullet}^{h1}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet\bullet}^{hL}}{n_{\bullet\bullet}^{h\bullet}} \right), h = 1, \dots, H \quad (3.3.7)$$

As massas destes pontos são:

$$(r_{\bullet}^h)_3 = \frac{n_{\bullet\bullet}^{h\bullet}}{n}, h = 1, \dots, H \quad (3.3.8)$$

Comparando os perfis linha das matrizes  $N$  e  $N_2$ , podemos observar que cada perfil linha de  $N_2$  é realmente o centróide dos respectivos grupos dos perfis linha de  $N$ . Tomemos, então, a definição de centróide dada em (2.3.7), e substituindo os perfis e suas massas como descrito acima, estabelecemos a seguinte relação

$$(R_{\bullet}^h)_2 = \frac{1}{(r_{\bullet}^h)_2} \sum_{i=1}^{I_h} (r_i^h)(R_i^h) \quad (3.3.9)$$

A veracidade desta relação pode ser verificada ao efetuarmos a multiplicação entre  $(r_{\bullet}^h)_2$  e  $(R_{\bullet}^h)_2$ , quando, então, obtemos

$$\left( \frac{n_{\bullet 1}^{h1}}{n}, \dots, \frac{n_{\bullet J_1}^{h1}}{n}, \dots, \frac{n_{\bullet 1}^{hL}}{n}, \dots, \frac{n_{\bullet J_L}^{hL}}{n} \right), h = 1, \dots, H. \quad (3.3.10)$$

Por outro lado, para cada  $i = 1, \dots, I_h$ , ao multiplicarmos  $r_i^h$  por  $R_i^h$ , obtemos

$$\left( \frac{n_{i1}^{h1}}{n}, \dots, \frac{n_{iJ_1}^{h1}}{n}, \dots, \frac{n_{i1}^{hL}}{n}, \dots, \frac{n_{iJ_L}^{hL}}{n} \right), h = 1, \dots, H.$$

Ao efetuar o somatório em  $i$ , obtemos a mesma expressão de (3.2.10).

Tomando os perfis linha de  $\mathbf{N}_1$  e  $\mathbf{N}_3$ , também, podemos mostrar que os perfis linha de  $\mathbf{N}_3$  são os centróides dos grupos de perfis linha de  $\mathbf{N}_1$ , isto é, que

$$(\mathbf{R}_{\bullet}^h)_3 = \frac{1}{(r_{\bullet}^h)_3} \sum_{i=1}^I (r_i^h)_1 (\mathbf{R}_i^h)_1 \quad (3.3.11)$$

Na matriz  $\mathbf{N}$  com estrutura de grupo, os perfis linha são representados por uma nuvem composta de várias outras nuvens, onde cada uma delas corresponde aos grupos desta matriz. A partir desta nuvem podemos obter três outras nuvens que representam os perfis linha das matrizes  $\mathbf{N}_1$ ,  $\mathbf{N}_2$  e  $\mathbf{N}_3$ .

Assim, da matriz  $\mathbf{N}$ , temos uma nuvem no espaço  $\mathbf{R}^J$ , formado por  $H$  nuvens compostas de  $I_h$  pontos. Tal nuvem pode ser expressa por

$$\mathbf{R} = [(\mathbf{R}_1^1)' \dots (\mathbf{R}_{I_1}^1)' \dots (\mathbf{R}_1^H)' \dots (\mathbf{R}_{I_H}^H)']' = \mathbf{D}_r^{-1} \mathbf{P} \quad (3.3.12)$$

No caso da matriz  $\mathbf{N}_1$ , como as colunas foram agrupadas, temos uma nuvem no espaço  $\mathbf{R}^L$ , que também é composta por  $I = \sum I_h$  pontos, formando  $H$  nuvens, e pode ser escrita como:

$$\mathbf{R}_1 = [(\mathbf{R}_1^1)_1' \dots (\mathbf{R}_{I_1}^1)_1' \dots (\mathbf{R}_1^H)_1' \dots (\mathbf{R}_{I_H}^H)_1']' = \mathbf{D}_{r_1}^{-1} \mathbf{P}_1 \quad (3.3.13)$$

No caso em que a matriz foi construída a partir do agrupamento de linhas, isto é, para  $\mathbf{N}_2$ , temos uma nuvem no espaço  $\mathbf{R}^J$  formada pelos  $H$  centróides dos grupos de  $\mathbf{N}$ , expressa por

$$\mathbf{R}_2 = [(\mathbf{R}_{\bullet}^1)_2' \dots (\mathbf{R}_{\bullet}^H)_2']' = \mathbf{D}_{r_2}^{-1} \mathbf{P}_2 \quad (3.3.14)$$

Quando agrupamos as linhas e as colunas, obtemos a matriz  $\mathbf{N}_3$ . Esta matriz está representada por uma nuvem no espaço  $\mathbf{R}^L$ . Esta nuvem é formada pelos  $H$  centróides dos grupos de  $\mathbf{N}_1$ , e é expressa por:

$$\mathbf{R}_3 = [(\mathbf{R}_{\bullet}^1)_3' \dots (\mathbf{R}_{\bullet}^H)_3']' = \mathbf{D}_{r_3}^{-1} \mathbf{P}_3. \quad (3.3.15)$$

Como cada perfil linha tem uma massa, cada uma destas nuvens de perfis linha terá um vetor de massa. Assim, os vetores de massa das nuvens  $\mathbf{R}$ ,  $\mathbf{R}_1$ ,  $\mathbf{R}_2$  e  $\mathbf{R}_3$  são, respectivamente,

$$\begin{aligned}
\mathbf{r} &= \left( \frac{n_{1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{I_1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{1\bullet}^{H\bullet}}{n}, \dots, \frac{n_{I_H\bullet}^{H\bullet}}{n} \right), \\
\mathbf{r}_1 &= \left( \frac{n_{1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{I_1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{1\bullet}^{H\bullet}}{n}, \dots, \frac{n_{I_H\bullet}^{H\bullet}}{n} \right), \\
\mathbf{r}_2 &= \left( \frac{n_{\bullet\bullet}^{1\bullet}}{n}, \dots, \frac{n_{\bullet\bullet}^{H\bullet}}{n} \right), \\
\mathbf{r}_3 &= \left( \frac{n_{\bullet\bullet}^{1\bullet}}{n}, \dots, \frac{n_{\bullet\bullet}^{H\bullet}}{n} \right).
\end{aligned} \tag{3.3.16}$$

Comparando os vetores de massas acima, percebemos que  $\mathbf{r} = \mathbf{r}_1$  e que  $\mathbf{r}_2 = \mathbf{r}_3$ . Estes resultados se justificam devido a dois fatos: Primeiro porque as nuvens  $\mathbf{R}$  e  $\mathbf{R}_1$  têm a mesma quantidade de pontos, o mesmo acontecendo com  $\mathbf{R}_2$  e  $\mathbf{R}_3$ . Em segundo lugar, a operação de somar as frequências das colunas, ao fazer o agrupamento, não altera a proporção das linhas. Como a massa foi definida como a proporção de indivíduos em cada linha em relação ao total de indivíduos da matriz, temos, portanto, que as massas dos pontos de  $\mathbf{R}$  são iguais às de  $\mathbf{R}_1$ , o mesmo acontecendo para  $\mathbf{R}_2$  e  $\mathbf{R}_3$ .

No capítulo 2, fizemos um comentário a respeito do fato de que, em análise de correspondência, o estudo das nuvens é feito tomando como base o seu centróide ou representante. Isto acontece porque ele, além de compor a diagonal do inverso da matriz que define a métrica em cada nuvem, é tomado como referência para a medida de dispersão da nuvem. Assim, definimos  $\mathbf{c}$ ,  $\mathbf{c}_1$ ,  $\mathbf{c}_2$  e  $\mathbf{c}_3$  como sendo os centróides das nuvens de perfis linha de  $N$ ,  $N_1$ ,  $N_2$  e  $N_3$ , dados respectivamente por:

$$\begin{aligned}
\mathbf{c} &= \left( \frac{n_{\bullet 1}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet j_1}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet 1}^{\bullet L}}{n}, \dots, \frac{n_{\bullet j_L}^{\bullet L}}{n} \right), \\
\mathbf{c}_1 &= \left( \frac{n_{\bullet\bullet}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet\bullet}^{\bullet L}}{n} \right), \\
\mathbf{c}_2 &= \left( \frac{n_{\bullet 1}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet j_1}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet 1}^{\bullet L}}{n}, \dots, \frac{n_{\bullet j_L}^{\bullet L}}{n} \right), \\
\mathbf{c}_3 &= \left( \frac{n_{\bullet\bullet}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet\bullet}^{\bullet L}}{n} \right).
\end{aligned} \tag{3.3.17}$$

Aqui, novos resultados aparecem ao se comparar o centróide das nuvens dos perfis linha das matrizes  $N$  e  $N_2$ , e os das matrizes  $N_1$  e  $N_3$ . Observa-se, então, que  $\mathbf{c} = \mathbf{c}_2$  e  $\mathbf{c}_1 = \mathbf{c}_3$ . Da forma como foi definido em (2.3.5), o centróide de uma nuvem é o perfil da linha formada pela soma de

todas as linhas da matriz. Deste modo, como as linhas de N<sub>2</sub> são agrupamentos das de N, a soma das linhas dessas duas matrizes são iguais. Além disso, os espaços vetoriais dos perfis linha dessas duas matrizes têm a mesma dimensão. A partir destes fatos, percebemos que o centróide da nuvem dos perfis linha de N é igual ao da nuvem dos perfis linha de N<sub>2</sub>. O mesmo argumento vale para os perfis linha de N<sub>1</sub> e N<sub>3</sub>.

No caso das colunas, os perfis são análogos e simétricos ao perfis linha (ver apêndice A) e suas relações também são análogas. Assim temos que os perfis coluna de N<sub>1</sub> (N<sub>3</sub>) são os centróides dos grupos dos perfis coluna de N (N<sub>2</sub>), estabelecendo as seguintes igualdades:

$$(C_{\bullet}^{\ell})_1 = \frac{1}{(c_{\bullet}^{\ell})_1} \sum_{j=1}^J (c_j^{\ell})(C_j^{\ell}) \quad (3.3.18)$$

e

$$(C_{\bullet}^{\ell})_3 = \frac{1}{(c_{\bullet}^{\ell})_3} \sum_{j=1}^J (c_j^{\ell})_2 (C_j^{\ell})_2. \quad (3.3.19)$$

As nuvens dos perfis coluna, seus vetores de massas e centróides também são similares, bem como as relações entre os vetores de massas e entre os centróides. Esta similaridade deve-se ao fato de as massas dos perfis linha serem os centróides dos perfis coluna e vice-versa.

Visto, então, as nuvens dos perfis linha e dos perfis coluna e suas relações, vamos apresentar as métricas e suas relações.

### 3.3.2 AS MÉTRICAS E SUAS RELAÇÕES

No caso em que a estrutura de grupo é considerada, também é a distância entre os pontos que nos mostra se existe ou não uma associação entre eles; contudo, agora, é possível estabelecermos algumas relações entre as estruturas métricas das diferentes nuvens. Para isso, vamos definir a estrutura métrica de cada uma das nuvem. Tais métricas são dadas por distância euclidiana generalizada, definida pelas matrizes diagonais onde, na diagonal, temos o inverso dos elementos dos seus respectivos centróides.

As métricas definidas para as nuvens dos perfis linha nos espaços  $\mathbf{R}^J$ ,  $\mathbf{R}^L$ ,  $\mathbf{R}^J$  e  $\mathbf{R}^L$ , referentes às matrizes  $N$ ,  $N_1$ ,  $N_2$  e  $N_3$ , são dadas, respectivamente, por

$$d^2((R_i^h), (R_{i'}^{h'})) = \sum_{\ell=1}^L \sum_{j=1}^{J_\ell} \left( \frac{n_{ij}^{h\ell}}{n_{i\bullet}^{h\bullet}} - \frac{n_{i'j}^{h'\ell}}{n_{i'\bullet}^{h'\bullet}} \right)^2 \frac{n_{\bullet\ell}}{n_{\bullet\bullet}^{\ell}} = ((R_i^h) - (R_{i'}^{h'}))' \mathbf{D}_c^{-1} ((R_i^h) - (R_{i'}^{h'}))$$

$$d^2((R_i^h)_1, (R_{i'}^{h'})_1) = \sum_{\ell=1}^L \left( \frac{n_{i\bullet}^{h\ell}}{n_{i\bullet}^{h\bullet}} - \frac{n_{i'\bullet}^{h'\ell}}{n_{i'\bullet}^{h'\bullet}} \right)^2 \frac{n_{\bullet\ell}}{n_{\bullet\bullet}^{\ell}} = ((R_i^h)_1 - (R_{i'}^{h'})_1)' \mathbf{D}_{c_1}^{-1} ((R_i^h)_1 - (R_{i'}^{h'})_1)$$

$$d^2((R_{\bullet}^h)_2, (R_{\bullet}^{h'})_2) = \sum_{\ell=1}^L \sum_{j=1}^{J_\ell} \left( \frac{n_{\bullet j}^{h\ell}}{n_{\bullet\bullet}^{h\bullet}} - \frac{n_{\bullet j}^{h'\ell}}{n_{\bullet\bullet}^{h'\bullet}} \right)^2 \frac{n_{\bullet\ell}}{n_{\bullet\bullet}^{\ell}} = ((R_{\bullet}^h)_2 - (R_{\bullet}^{h'})_2)' \mathbf{D}_{c_2}^{-1} ((R_{\bullet}^h)_2 - (R_{\bullet}^{h'})_2)$$

$$d^2((R_{\bullet}^h)_3, (R_{\bullet}^{h'})_3) = \sum_{\ell=1}^L \left( \frac{n_{\bullet\bullet}^{h\ell}}{n_{\bullet\bullet}^{h\bullet}} - \frac{n_{\bullet\bullet}^{h'\ell}}{n_{\bullet\bullet}^{h'\bullet}} \right)^2 \frac{n_{\bullet\ell}}{n_{\bullet\bullet}^{\ell}} = ((R_{\bullet}^h)_3 - (R_{\bullet}^{h'})_3)' \mathbf{D}_{c_3}^{-1} ((R_{\bullet}^h)_3 - (R_{\bullet}^{h'})_3)$$

As métricas definidas para os perfis coluna nos espaços  $\mathbf{R}^I$ ,  $\mathbf{R}^H$ ,  $\mathbf{R}^I$  e  $\mathbf{R}^H$ , referentes às matrizes  $N$ ,  $N_1$ ,  $N_2$  e  $N_3$  são análogas às dos perfis linha.

Para estudar as relações existentes entre estas estruturas métricas, lembremo-nos primeiramente que, no item anterior, ao definirmos os centróides, foram estabelecidas as seguintes relações:  $\mathbf{c} = \mathbf{c}_2$ ,  $\mathbf{c}_1 = \mathbf{c}_3$ ,  $\mathbf{r} = \mathbf{r}_1$  e  $\mathbf{r}_2 = \mathbf{r}_3$ . E também que as matrizes diagonais que definem as distâncias têm como elementos o inverso dos elementos de seus respectivos centróides. Assim, percebe-se que nuvens que têm centróides iguais têm métricas iguais. Portanto as métricas definidas em  $\mathbf{R}$  e  $\mathbf{R}_2$ , são iguais e estas nuvens estão no mesmo espaço vetorial  $\mathbf{R}^J$ . O mesmo acontece com as métricas definidas em  $\mathbf{R}_1$  e  $\mathbf{R}_3$ , que estão no espaço vetorial  $\mathbf{R}^L$ . Por sua vez, as métricas definidas em  $\mathbf{C}$  e  $\mathbf{C}_1$  são iguais e estas nuvens estão no mesmo espaço  $\mathbf{R}^I$ . O mesmo acontece com  $\mathbf{C}_2$  e  $\mathbf{C}_3$  no espaço  $\mathbf{R}^H$ .

Para estudar as diferenças entre os grupos, tomamos os seus centróides como seus representantes e com eles calculamos os eixos direcionais. Mas, estes centróides são os pontos da nuvem  $\mathbf{R}_2$ , enquanto que todos os outros pontos da nuvem são os pontos de  $\mathbf{R}$ . Ora, vimos acima que  $\mathbf{R}$  e  $\mathbf{R}_2$  estão no mesmo espaço e têm a mesma métrica. Nuvens com métricas iguais possibilitam a representação dos pontos de uma delas como pontos suplementares na outra. Assim, podemos expressar os pontos da nuvem  $\mathbf{R}$  como pontos suplementares, junto com os pontos de  $\mathbf{R}_2$ . Da mesma forma, podemos expressar os pontos de  $\mathbf{R}_1$ ,  $\mathbf{C}$  e  $\mathbf{C}_2$  como pontos suplementares em  $\mathbf{R}_3$ ,  $\mathbf{C}_1$  e  $\mathbf{C}_3$ .

Antes de calculamos os eixos direcionais, vamos discutir as relações entre as inércias totais das nuvens.

### 3.3.2.1 RELAÇÕES ENTRE AS INÉRCIAS TOTAIS DAS NUUVENS

As inércias das nuvens  $\mathbf{R}$  e  $\mathbf{R}_2$ ,  $\mathbf{R}_1$  e  $\mathbf{R}_3$ ,  $\mathbf{C}$  e  $\mathbf{C}_1$  e  $\mathbf{C}_2$  e  $\mathbf{C}_3$  são comparáveis, já que cada duas delas estão no mesmo espaço vetorial e têm definida a mesma métrica. É de se esperar que a inércia de  $\mathbf{R}$  seja maior que a de  $\mathbf{R}_2$ , o mesmo acontecendo com  $\mathbf{R}_1$  e  $\mathbf{R}_3$ , com  $\mathbf{C}$  e  $\mathbf{C}_1$  e com  $\mathbf{C}_2$  e  $\mathbf{C}_3$ , já que as nuvens  $\mathbf{R}_2$ ,  $\mathbf{R}_3$ ,  $\mathbf{C}_1$  e  $\mathbf{C}_3$  surgem da condensação dos grupos das nuvens  $\mathbf{R}$ ,  $\mathbf{R}_1$ ,  $\mathbf{C}$  e  $\mathbf{C}_2$ , respectivamente. De fato, com respeito a qualquer subespaço, temos o resultado clássico de que a inércia total de uma nuvem é igual à soma das inércias entre os grupos mais a soma das inércias dentro do grupo. Vamos enunciar mais formalmente e demonstrar este resultado apenas para o caso das nuvens  $\mathbf{R}$  e  $\mathbf{R}_2$ .

Sejam  $\mathbf{R}^h$ ,  $h=1, \dots, H$ , as nuvens compostas pelos pontos  $\mathbf{R}_i^1$ ,  $i=1, \dots, I_1$ ;  $\mathbf{R}_i^2$ ,  $i=1, \dots, I_2$ ; ... ;  $\mathbf{R}_i^H$ ,  $i=1, \dots, I_H$ , com as respectivas massas  $r_i^h$ ,  $i=1, \dots, I_h$  e  $h=1, \dots, H$ .  $\mathbf{R}$  é a nuvem dos  $I = \sum_{h=1}^H I_h$  perfis linha no espaço euclidiano  $J$ -dimensional ponderado e estruturado pela métrica definida pela matriz simétrica positiva-definida  $D_c^{-1}$ . E  $\mathbf{R}^h$ ,  $h=1, \dots, H$ , são as  $H$  nuvens que formam a nuvem  $\mathbf{R}$ .

Sejam  $\mathbf{R}_\bullet^h = \frac{1}{r_\bullet^h} \sum_{i=1}^{I_h} r_i^h \mathbf{R}_i^h$ ,  $h=1, \dots, H$ , e  $\mathbf{c} = \frac{1}{\bar{r}} \sum_{h=1}^H r_\bullet^h \mathbf{R}_\bullet^h$  os centróides das nuvens  $\mathbf{R}^h$  e o centróide geral da nuvem  $\mathbf{R}$ , respectivamente, onde  $r_\bullet^h = \sum_{i=1}^{I_h} r_i^h$  são as massas dos centróides das nuvens  $\mathbf{R}^h$  e  $\bar{r} = \sum_{h=1}^H r_\bullet^h$  a massa do centróide da nuvem  $\mathbf{R}$ . Vamos mostrar que a inércia total da nuvem  $\mathbf{R}$ , com respeito ao centróide geral  $\mathbf{c}$ , é igual a soma das inércias totais das  $H$  nuvens  $\mathbf{R}^h$ , com respeito aos seus respectivos centróides  $\mathbf{R}_\bullet^h$ , mais a soma das inércias individuais dos centróides  $\mathbf{R}_\bullet^h$ , em relação ao centróide geral  $\mathbf{c}$ .

Conforme a notação da inércia, definida a partir de (2.3.11), podemos escrever a afirmação acima como sendo a seguinte igualdade:

$$In_{\mathbf{c}}(\mathbf{R}) = \sum_{h=1}^H In_{\mathbf{R}_\bullet^h}(\mathbf{R}^h) + \sum_{h=1}^H In_{\mathbf{c}}(\mathbf{R}_\bullet^h) \quad (3.3.21)$$

Para demonstrar esta igualdade, comecemos observando a expressão da inércia dada em (2.3.12). Esta expressão nos diz que a inércia total de uma nuvem é formada pela soma das inércias individuais dos pontos dessa nuvem. Inércias estas calculadas em relação a um mesmo ponto. Assim, para a nuvem  $\mathbf{R}$ , temos

$$In_c(\mathbf{R}) = \sum_{h=1}^H \sum_{i=1}^{I_h} In_c(\mathbf{R}_i^h). \quad (3.3.22)$$

Mas como  $\sum_{i=1}^{I_h} In_c(\mathbf{R}_i^h) = In_c(\mathbf{R}^h)$  isto é, a soma das inércias individuais dos  $I_h$  pontos da  $h$ -ésima nuvem é a inércia total desta nuvem, então podemos considerar o segundo membro da expressão acima como a soma das inércias totais de cada uma das  $H$  nuvens, em relação ao centróide geral  $c$ . Assim, temos que

$$In_c(\mathbf{R}) = \sum_{h=1}^H In_c(\mathbf{R}^h). \quad (3.3.23)$$

Mas, neste caso, as inércias totais das nuvens  $\mathbf{R}^h$  não estão calculadas em relação aos seus centróides, e sim em relação ao centróide geral. Para expressar a relação entre as inércias calculadas entre estes dois pontos, tomemos o Teorema de Huyghens dado pela expressão (2.3.15) que reescrevemos a seguir

$$In_y(\mathbf{R}) = In_{\bar{R}}(\mathbf{R}) + In_y(\bar{\mathbf{R}}).$$

Esta expressão nos dá a relação entre a inércia de cada nuvem, calculada em relação ao seu respectivo centróide, e a inércia destas nuvens, calculada em relação a um ponto qualquer. Ao utilizar este resultado para as nuvens  $\mathbf{R}^h$ ,  $h=1, \dots, H$ , o ponto qualquer  $y$  é substituído pelo centróide geral  $c$  e  $\bar{\mathbf{R}}$  pelos centróides  $\mathbf{R}_\bullet^h$  de cada nuvem. Estas expressões têm a seguinte forma

$$In_c(\mathbf{R}^h) = In_{\mathbf{R}_\bullet^h}(\mathbf{R}^h) + In_c(\mathbf{R}_\bullet^h). \quad (3.3.24)$$

Assim, substituindo esta relação, dada em (3.3.24), na expressão da inércia total de  $\mathbf{R}$ , dada em (3.3.23), e desenvolvendo o somatório em  $h$ , temos o que queríamos demonstrar, isto é, temos que

$$In_c(\mathbf{R}) = \sum_{h=1}^H In_{\mathbf{R}_\bullet^h}(\mathbf{R}^h) + \sum_{h=1}^H In_c(\mathbf{R}_\bullet^h).$$

A nuvem formada pelos centróides  $\mathbf{R}_\bullet^h$ ,  $h=1, \dots, H$ , é a que representamos por  $\mathbf{R}_2$ , formada pelos perfis das linhas de  $N_2$ . Portanto, esse resultado nos mostra a relação existente entre as inércias da matrizes formadas pelos perfis linha de  $N$  e  $N_2$  e expressa por

$$In_{\mathbf{c}}(\mathbf{R}) = \sum_{h=1}^H In_{\mathbf{R}_h^*}(\mathbf{R}^h) + In_{\mathbf{c}}(\mathbf{R}_2). \quad (3.3.25)$$

Isto nos mostra que a soma das inércias dos centróides é também sempre menor que a soma das inércias individuais, por um decréscimo igual à soma das inércias dentro dos grupos. Como  $\mathbf{c}$ , o centróide da nuvem dos perfis linha de  $\mathbf{N}$ , é igual a  $\mathbf{c}_2$ , o centróide da nuvem dos perfis linha de  $\mathbf{N}_2$ , podemos dizer que a inércia total da nuvem  $\mathbf{R}$ , em relação ao seu centróide, é maior ou igual à inércia total da nuvem  $\mathbf{R}_2$ , em relação ao seu centróide, isto é,

$$In_{\mathbf{c}}(\mathbf{R}) \geq In_{\mathbf{c}_2}(\mathbf{R}_2). \quad (3.3.26)$$

Resultados análogos a estes, são válidos:

para o caso das nuvens  $\mathbf{R}_1$  e  $\mathbf{R}_3$ , isto é,

$$In_{\mathbf{c}_1}(\mathbf{R}_1) = \sum_{h=1}^H In_{\mathbf{R}_h^*}(\mathbf{R}_1^h) + In_{\mathbf{c}_3}(\mathbf{R}_3) \quad (3.3.27)$$

$$In_{\mathbf{c}_1}(\mathbf{R}_1) \geq In_{\mathbf{c}_3}(\mathbf{R}_3) \quad (3.3.28)$$

para os casos das nuvens  $\mathbf{C}$  e  $\mathbf{C}_1$ , isto é,

$$In_{\mathbf{r}}(\mathbf{C}) = \sum_{\ell=1}^L In_{\mathbf{C}_\ell^*}(\mathbf{C}^\ell) + In_{\mathbf{r}_1}(\mathbf{C}_1) \quad (3.3.29)$$

$$In_{\mathbf{r}}(\mathbf{C}) \geq In_{\mathbf{r}_1}(\mathbf{C}_1) \quad (3.3.30)$$

e para os casos das nuvens  $\mathbf{C}_2$  e  $\mathbf{C}_3$ , isto é,

$$In_{\mathbf{r}_2}(\mathbf{C}_2) = \sum_{\ell=1}^L In_{\mathbf{C}_\ell^*}(\mathbf{C}_2^\ell) + In_{\mathbf{r}_3}(\mathbf{C}_3) \quad (3.3.31)$$

$$In_{\mathbf{r}_2}(\mathbf{C}_2) \geq In_{\mathbf{r}_3}(\mathbf{C}_3). \quad (3.3.32)$$

Desta forma, isto nos mostra que a condensação dos grupos em pontos reduz a inércia total de todas as nuvens.

Estabelecidas estas relações, vamos discutir o ajuste das nuvens e suas relações

### 3.3.3 OS AJUSTES DAS NUVENS E SUAS RELAÇÕES

Nós havíamos colocado, no capítulo 2, que ajustar uma nuvem a um subespaço significa encontrar um subespaço na qual os vetores de sua base são as direções em que os pontos, nela projetados, têm as maiores inércias. Isto foi feito para cada direção, através da expressão (2.3.24) que reescrevemos a seguir:

$$(D_r^{-1/2} Xv)' D_r (D_r^{-1/2} Xv) = v' X' X v = v' S v$$

Agora estamos considerando a estrutura de grupo e estamos interessados em estudar as diferenças entre eles. Por isto, interessa-nos maximizar a inércia, não mais de todos os pontos da nuvem, mas apenas dos representantes de cada grupo. Desta forma, tomemos apenas as matrizes que têm os centróides como elementos; ou seja, maximizaremos as expressões das matrizes que definem o ajuste das nuvens  $R_2$ ,  $R_3$ ,  $C_1$  e  $C_3$ .

A matriz que define o ajuste da nuvem  $R_2$  no espaço  $R^J$ , é semelhante a matriz  $S$  definida em (2.3.21), e é a seguinte:

$$\begin{aligned} S_2 &= (P_2 - r_2 c_2') D_{r_2}^{-1} (P_2 - r_2 c_2')' D_{c_2}^{-1} \\ S_2 &= X_2' X_2 \text{ onde} \\ X_2 &= D_{r_2}^{-1/2} (P_2 - r_2 c_2') D_{c_2}^{-1/2}. \end{aligned} \quad (3.3.33)$$

Por sua vez, a matriz que define o ajuste da nuvem  $R_3$  no espaço  $R^L$  é dada por:

$$\begin{aligned} S_3 &= (P_3 - r_3 c_3') D_{r_3}^{-1} (P_3 - r_3 c_3')' D_{c_3}^{-1} \\ S_3 &= X_3' X_3 \text{ onde} \\ X_3 &= D_{r_3}^{-1/2} (P_3 - r_3 c_3') D_{c_3}^{-1/2}, \end{aligned} \quad (3.3.34)$$

a matriz que define o ajuste da nuvem  $C_1$  no espaço  $R^I$  é dada por

$$\begin{aligned} Z_1 &= (P_1' - c_1 r_1')' D_{c_1}^{-1} (P_1' - c_1 r_1') D_{r_1}^{-1} \\ Z_1 &= Y_1' Y_1 \text{ onde} \\ Y_1 &= D_{c_1}^{-1/2} (P_1' - c_1 r_1')' D_{r_1}^{-1/2} \end{aligned} \quad (3.3.35)$$

e a matriz que define o ajuste da nuvem  $C_3$  no espaço  $R^H$  é dada por

$$\begin{aligned} Z_3 &= (P_3' - c_3 r_3')' D_{c_3}^{-1} (P_3' - c_3 r_3') D_{r_3}^{-1} \\ Z_3 &= Y_3' Y_3 \text{ onde} \\ Y_3 &= D_{c_3}^{-1/2} (P_3' - c_3 r_3')' D_{r_3}^{-1/2}. \end{aligned} \quad (3.3.36)$$

A partir da decomposição em valores singulares das matrizes que as compõem, obtém-se os vetores direcionais de cada nuvem de centróides e, portanto, as coordenadas desses ponto (ver apêndice A).

Para analisar os pontos de  $R$ , usaremos a estrutura de  $R_2$ , colocando os pontos de  $R$  como pontos suplementares. Da mesma forma, para a análise dos pontos de  $R_1$ ,  $C$  e  $C_2$  usaremos as estruturas de  $R_3$ ,  $C_1$  e  $C_3$ , respectivamente, colocando os seus pontos como pontos suplementares.

Ao fazê-lo, estamos maximizando a inércia entre e mantendo a inércia dentro dos grupos, o que é equivalente a fazer a Análise Discriminante dos grupos. Desta forma, percebe-se a veracidade da afirmação de Greenacre a respeito das relações entre estes dois tipos de análise. Ele coloca que “a Análise de Correspondência em  $N_1$ ,  $N_2$  e em  $N_3$  é efetivamente uma Análise Discriminante” e que “a Análise de Correspondência em  $N$  também pode ser vista como uma Análise Discriminante se consideramos cada linha como um grupo, onde os grupos não tem variabilidade dentro”.

Os resultados da análise de correspondência de cada uma das quatro matrizes, obtidas a partir de uma tabela de contingência com estrutura de grupo, é dada separadamente no apêndice A. Agora veremos como as inércias principais dos eixos das nuvens de centróides estão relacionadas com as dos eixos das nuvens de perfis.

### 3.3.3.1 RELAÇÕES ENTRE AS INÉRCIAS DOS EIXOS

Vimos, na seção 3.3.2.1, que a condensação dos grupos em pontos reduz a inércia total da nuvem. Agora vamos mostrar um outro resultado, que é uma consequência desse, e nos mostra a redução da inércia, agora nos eixos principais definidos pelo ajuste. Na investigação das análises de  $R$  e  $R_1$ ,  $R_2$  e  $R_3$ ,  $C$  e  $C_1$ ,  $C_2$  e  $C_3$  é sempre interessante comparar as inércias de seus eixos principais. Em termos das inércias principais, temos que, em qualquer subespaço, as inércias principais dos eixos das nuvens agrupadas é sempre menor que as inércias principais dos eixos das nuvens não agrupadas. A seguir, enunciaremos mais formalmente este resultado para  $R$  e  $R_2$ .

Sejam  $R_i^h$ ,  $i=1,\dots,J_h$  e  $h=1,\dots,H$  e  $R_{\bullet}^h$ ,  $h=1,\dots,H$ , os pontos das nuvens  $R$  e  $R_2$ , respectivamente, com as respectivas massas  $r_i^h$  e  $r_{\bullet}^h$ . Estes pontos estão em um espaço euclidiano  $J$ -dimensional ponderado e estruturado por uma métrica definida pela matriz simétrica positiva-definida  $D_c^{-1}$ . Sejam  $v_1, v_2, \dots, v_K$  e  $v_1', v_2', \dots, v_{K'}$  os eixos que nos dão as dimensões do espaço vetorial em que se encontra a nuvem  $R$  e  $R_2$ , respectivamente. Sejam ainda  $\lambda_1, \lambda_2, \dots, \lambda_K$  e  $\lambda'_1, \lambda'_2, \dots, \lambda'_{K'}$  as inércias principais dos eixos da nuvem  $R$  e  $R_2$ . Essas inércias principais dadas em termos das inércias dos perfis, são as somas das projeções das inércias de todos os perfis linha sobre cada eixo. Assim temos que

$$\lambda_k = \sum_{h=1}^H \sum_{i=1}^{I_h} [In_c(\mathbf{R}_i^h)]_k, \quad k=1, \dots, K \quad (3.3.37)$$

onde  $[In_c(\mathbf{R}_i^h)]_k$  está representando a componente da inércia dos pontos  $\mathbf{R}_i^h$ ,  $i=1, \dots, I_h$  e  $h=1, \dots, H$  sobre o eixo  $v_k$ . Lembramos que as inércias principais são dadas pelos autovalores associados aos autovetores da matriz que define o ajuste do subespaço ótimo e que os autovetores nos dão as direções desses eixos. Desta forma, devemos mostrar que as inércias principais  $\lambda_1, \lambda_2, \dots$  dos eixos da nuvem dos perfis linha de  $\mathbf{R}$  são maiores ou iguais às respectivas inércias principais  $\lambda'_1, \lambda'_2, \dots$  dos eixos da nuvem dos perfis linha de  $\mathbf{R}_2$ , isto é, devemos mostrar que

$$\lambda_1 \geq \lambda'_1, \lambda_2 \geq \lambda'_2, \dots$$

Para demonstrar este resultado, observe que  $K' \leq K$ , já que estamos supondo que a dimensionalidade dos  $I$  pontos de  $\mathbf{R}$  seja  $K$  e a dos  $H$  centróides de  $\mathbf{R}_2$  seja  $K'$ . O resultado é trivialmente verdadeiro para  $k > K'$ , tendo neste caso  $\lambda'_k = 0$ . Para  $k \leq K'$ , considere os seguintes dois subespaços do espaço  $K$ -dimensional, onde estão representados todos os perfis linha. Este é o subespaço que denominamos de Subespaço Ajustado. O primeiro subespaço  $\mathbf{A}$  é formado pelos eixos  $v_t$ , que tenham inércias principais  $\lambda_t = [\sum_{h=1}^H \sum_{i=1}^{I_h} In_c(\mathbf{R}_i^h)]_{v_t}$ ,  $t=1, \dots, k$  menor ou igual a  $\lambda_k = [In_c(\mathbf{R})]_{v_k}$ , isto é,

$$\mathbf{A} = \{ v_t \text{ tal que } \lambda_k \geq [In_c(\mathbf{R})]_{v_t}, \text{ para } t=1, \dots, k \}. \quad (3.3.38)$$

O segundo subespaço  $\mathbf{B}$  é formado pelos eixos  $v_t'$ , que tenham inércias principais  $\lambda'_t =$

$$[\sum_{h=1}^H In_c(\mathbf{R}_i^h)]_{v_t'} = [In_c(\mathbf{R}_2)]_{v_t'}, \text{ maiores ou iguais a } \lambda'_k.$$

$$\mathbf{B} = \{ v_t', \text{ tal que } [In_c(\mathbf{R}_2)]_{v_t'} \geq \lambda_k, t=k, \dots, K' \}. \quad (3.3.39)$$

O primeiro subespaço exclui os primeiros  $k-1$  eixos principais dos  $I$  perfis linha e tem dimensão  $K - (k-1) = K - k + 1$  enquanto o segundo subespaço inclui os  $k$  eixos principais dos  $H$  centróides e sua dimensão é  $k$ . Assim a soma dessas dimensões é  $K - k + 1 + k = K + 1$  e, portanto, a interseção desses dois subespaços deve ter dimensão maior ou igual a 1 (um). Podemos assim assumir a existência de um eixo  $\mathbf{u}$  que é comum aos dois subespaços. O resultado da seção 3.3.2.1, dado em (3.3.21), é aplicável às componentes de todos os perfis linha de  $\mathbf{R}$  e de  $\mathbf{R}_2$  em relação a esse eixo, e,

portanto, a soma das componentes das inércia de todos os perfis linha da nuvem  $\mathbf{R}$  sobre o eixo  $\mathbf{u}$  é maior ou igual à soma das projeções das inércias de todos os perfis linha da nuvem  $\mathbf{R}_2$  dos centróides, isto é

$$[In_c(\mathbf{R})]_{\mathbf{u}} \geq [In_c(\mathbf{R}_2)]_{\mathbf{u}}. \quad (3.3.40)$$

Por definição dos dois subespaços aos quais  $\mathbf{u}$  pertence temos

$$\lambda_{k.} \geq [In_c(\mathbf{R})]_{\mathbf{u}} \text{ e } [In_c(\mathbf{R}_2)]_{\mathbf{u}} \geq \lambda_k^2$$

o que implica

$$\lambda_{k.} \geq \lambda_k', k=1, \dots, K. \quad (3.3.41)$$

Isto nos mostra que o processo de condensação dos pontos dentro dos grupos reduz o momento de inércia em todas as direções.

Visto, então, as relações entre as varias matrizes proveniente de uma tabela de contingência com estrutura de grupos, passamos à aplicação.

## 4. APLICAÇÃO DA ANÁLISE DE CORRESPONDÊNCIA EM GRUPOS

### 4.1 INTRODUÇÃO

Esta aplicação faz parte de um projeto maior que vem sendo desenvolvido sob a coordenação da pesquisadora Dra. Maria Coleta F. A. de Oliveira, do Núcleo de Estudos de População da Universidade Estadual de Campinas (NEPO/UNICAMP) e que tem por objetivo conhecer as circunstâncias em que vivem as pessoas de famílias cujo núcleo não é o casal. Nos dados desta aplicação são observadas partições na população e o interesse de análise está na comparação das características da população dentro e entre as classes desta partição. Assim, a análise estatística destes dados recaiu na utilização da Análise de Correspondência em Grupos. Neste capítulo, apresentamos a obtenção dos resultados de forma detalhada, visando facilitar a compreensão da correspondência entre a teoria e a prática.

Para isso, na seção 4.2, apresentamos os dados a serem analisados, descrevendo as variáveis e a população em estudo. Na seção 4.3, apresentamos os resultados da análise de correspondência, como exposto no capítulos 2 e 3. E na seção 4.4, apresentamos os gráficos, a descrição dos pontos representantes de cada grupo e a dos pontos suplementares. Nesta fase apresentamos várias observações, muitas delas creditadas à Dra. Maria Coleta F. A. de Oliveira, que por serem preliminares e obtidas por comunicação oral, devo enfatizar que nenhuma delas necessariamente estarão de acordo com a forma ou conteúdo final das análises, a serem publicadas conjuntamente.

## 4.2 OS DADOS

Os dados analisados neste trabalho foram obtidos a partir dos censos demográficos de 1970, 1980 e 1991. Inicialmente, a população em estudo era a de homens sem companheiras no Estado de São Paulo e o objetivo do estudo, verificar a influência de variáveis descritoras na localização doméstica dos indivíduos da população. Desta forma, foi tomada como variável resposta a *Localização Doméstica* dentro de um domicílio onde o indivíduo em estudo não tem cônjuge. Esta variável caracteriza as alternativas de organização do domicílio dos homens sem companheiras. Suas categorias foram construídas com base na composição do domicílio, segundo as relações de parentesco entre seus componentes. Assim, a categoria em que foram agrupados os indivíduos que *vivem em grupos de pessoas sem laços de parentesco*, chamamos de **Localização 1 (L1)**; a dos que *vivem sós*, chamamos **Localização 2 (L2)**; a categoria dos que *vivem como chefe de família monoparental*<sup>(2)</sup>, chamamos **Localização 3 (L3)**; a dos que *vivem em grupos aparentados*, chamamos **Localização 4 (L4)** e a dos indivíduos que *vivem em família chefiada por outro*, chamamos de **Localização 5 (L5)**.

Para uma melhor compreensão da variável resposta, foram consideradas algumas características dos indivíduos desta população. Uma destas características é a idade, que foi categorizada em: indivíduos *jovens* de 15(quinze) a 24(vinte e quatro) anos; indivíduos *maduros* de 25 (vinte e cinco) a 59 (cinquenta e nove) anos e indivíduos *idosos* aqueles com 60 (sessenta) anos ou mais. Uma outra característica de interesse é a situação conjugal, que divide a população em indivíduos *solteiros*, *separados ou divorciados* e *viúvos*. Em um estudo preliminar observou-se uma forte associação entre as categorias jovens e solteiros, maduros e separados ou divorciados e idosos e viúvos. A categoria maduros foi originada da junção das categorias dos indivíduos de 25 a 39 anos e dos de 40 a 59 anos que neste mesmo estudo observou-se uma homogeneidade nos resultados destas categorias. Conforme relatório técnico entregue pela Dra Maria Coleta F. A. de Oliveira à PAEP (Pinho, A. et all 1997).

Assim, foi definida uma nova variável denominada de *Etapas do Ciclo de Vida*. Suas categorias são denominadas de: **1ª etapa de vida (E1)** para os indivíduos *solteiros de 15 a 24 anos*; **2ª etapa de vida (E2)** para os *separados ou divorciados de 25 a 59 anos* e **3ª etapa de**

---

<sup>(2)</sup> Família Monoparental é aquela constituída por apenas um dos conjuges e os filhos.

**vida (E3)** para os *viúvos de 60 anos ou mais*.

Desta forma a população em estudo passou a ser apenas os homens solteiros de 15 a 24 anos, os separados ou divorciados de 25 a 59 anos e os viúvos de 60 anos ou mais, no Estado de São Paulo. E a variável resposta Localização Doméstica será estudada mantendo controle das etapas de vida, que geram, portanto, as partições da população nas categorias da variável Localização Doméstica em cada censo.

As variáveis descritoras consideradas neste estudo são a **Instrução**, a **Renda Familiar per Capita** e a **Localização de Domicílio**, pois foram levantadas como possíveis explanatórias da variável Localização Doméstica.

A Instrução é categorizada em anos de estudo, considerando também os dados onde os indivíduos não declararam sua instrução. Assim, **Instrução não declarada (In)** é o nome da categoria em que foram agrupados os *indivíduos que não declararam instrução*. Chamamos: de **Instrução 0 (I0)** a categoria em que os *indivíduos não têm instrução*; de **Instrução 1 (I1)** a categoria em que os *indivíduos têm de 1 a 4 anos de estudo*; de **Instrução 2 (I2)** a categoria em que os *indivíduos têm de 5 a 8 anos de estudo*; de **Instrução 3 (I3)** a categoria em que os *indivíduos têm de 9 a 11 anos de estudo* e de **Instrução 4 (I4)** a categoria em que os *indivíduos têm 12 ou mais anos de estudo*.

As categorias da renda, por sua vez, foram definidas a partir da divisão da população em decis de renda<sup>(3)</sup>. Temos assim 12 (doze) categorias, dez das quais -as chamadas de **Renda 1 (R1)** até **Renda 10 (R10)**- são compostas pelos *indivíduos do primeiro ao décimo decil*. As outras são aquelas composta pelos *indivíduos sem rendimentos*, chamada **Renda 0 (R0)** e a dos que *não declararam rendimento*, chamada **Renda não declarada (Rn)**.

A localização de domicílio foi categorizada em apenas **Zona urbana (Zu)** e **Zona rural (Zr)**.

Desta forma, o problema permite considerar a variável Localização Doméstica -controlada pelas etapas de vida- como a variável a ser descrita, variável dependente ou resposta, enquanto que as outras são variáveis descritoras, independentes e 'candidatas' a explanatórias. Então, a tabulação cruzada da variável Localização Doméstica com cada uma das outras é de interesse, e além disso é possível enfocar essas relações de forma reduzida a uma única análise.

<sup>(3)</sup> Mais detalhes sobre estas categorias pode ser encontrada em Pinho (1997)



Para isso, os dados a analisar são apresentados através da representação, em tabelas de dupla entrada, determinada pela especificidade do problema:

- as linhas da tabela estão estruturadas conforme as categorias da variável Localização Doméstica nas etapas de vida, para os três censos considerados.

- as colunas estão estruturadas pela justaposição das categorias de cada uma das variáveis descritoras consideradas. Esta justaposição se justifica, pois temos o interesse em analisar a influência de cada uma das variáveis descritoras na presença das outras.

A tabela 5, denominada de *Tabela de Dados*, apresentada na página seguinte, é, então, construída pela concatenação, por coluna, de três tabelas correspondentes ao cruzamento da variável Localização Doméstica, controlada pelas etapas de vida, com cada uma das descritoras (Instrução, Renda e Localização de Domicílio). E da concatenação, na estrutura das linhas, de cada uma das tabelas acima descritas, por ano dos censos.

Desta forma, a tabela 5, comparada à tabela 4, tem a estrutura de grupo apenas nas linhas, determinada pelo interesse em examinar o comportamento da Localização Doméstica, associado a cada descritora, simultaneamente, nos três censos.

TABELA 5: Tabela de dados

|       |    | INSTRUÇÃO |         |         |         |        |       | RENDIMENTO |        |        |        |        |        |
|-------|----|-----------|---------|---------|---------|--------|-------|------------|--------|--------|--------|--------|--------|
|       |    |           |         |         |         |        |       |            |        |        |        |        |        |
|       |    | E1        | E2      | E3      | E4      | E5     | E6    | R1         | R2     | R3     | R4     | R5     |        |
| 1     | E1 | 781       | 1922    | 447     | 924     | 3413   | 0     | 1382       | 117    | 0      | 420    | 0      | 0      |
|       | E2 | 513       | 637     | 0       | 0       | 18     | 0     | 0          | 0      | 0      | 200    | 138    | 0      |
|       | E3 | 111       | 419     | 0       | 0       | 100    | 0     | 0          | 0      | 0      | 0      | 0      | 117    |
|       | L1 | 1378      | 2377    | 447     | 924     | 3731   | 0     | 1382       | 117    | 0      | 620    | 138    | 117    |
| 2     | E1 | 5111      | 7194    | 2399    | 1108    | 1499   | 0     | 392        | 0      | 379    | 305    | 399    | 498    |
|       | E2 | 3394      | 2179    | 498     | 1108    | 1300   | 0     | 194        | 0      | 305    | 599    | 139    | 1399   |
|       | E3 | 4014      | 1306    | 110     | 101     | 299    | 0     | 1308       | 112    | 399    | 202    | 396    | 1113   |
|       | L2 | 12519     | 10377   | 2997    | 2314    | 3698   | 0     | 1694       | 112    | 1083   | 1105   | 1133   | 3107   |
| 3     | E1 | 411       | 1115    | 299     | 0       | 297    | 0     | 0          | 403    | 495    | 99     | 109    | 137    |
|       | E2 | 1911      | 8111    | 399     | 1111    | 700    | 0     | 99         | 713    | 1113   | 1304   | 1311   | 399    |
|       | E3 | 9611      | 11110   | 197     | 399     | 491    | 100   | 137        | 496    | 1197   | 1401   | 1303   | 1091   |
|       | L3 | 10974     | 10416   | 696     | 1101    | 1398   | 100   | 396        | 1402   | 1415   | 1704   | 3203   | 3196   |
| 4     | E1 | 6142      | 10698   | 3953    | 3644    | 1959   | 411   | 119        | 1068   | 1208   | 3094   | 4099   | 4011   |
|       | E2 | 1039      | 1719    | 310     | 397     | 218    | 0     | 0          | 94     | 397    | 410    | 491    | 334    |
|       | E3 | 1094      | 1335    | 135     | 100     | 135    | 0     | 99         | 99     | 371    | 412    | 102    | 402    |
|       | L4 | 11274     | 13332   | 4398    | 4131    | 3112   | 411   | 318        | 1760   | 3676   | 3906   | 4892   | 4967   |
| 5     | E1 | 112738    | 74747   | 199417  | 184195  | 81090  | 630   | 5987       | 138499 | 143990 | 151355 | 139544 | 138316 |
|       | E2 | 6811      | 11789   | 699     | 1196    | 390    | 110   | 109        | 713    | 1113   | 1309   | 1371   | 1308   |
|       | E3 | 10911     | 11391   | 35      | 366     | 1323   | 0     | 134        | 4940   | 4995   | 3672   | 3199   | 1966   |
|       | L5 | 130659    | 87817   | 200647  | 186687  | 81090  | 740   | 6208       | 141192 | 149195 | 166616 | 143004 | 141287 |
| 6     | E1 | 3111      | 11411   | 4955    | 4949    | 1734   | 64    | 11         | 111    | 30     | 343    | 985    | 1144   |
|       | E2 | 477       | 1311    | 491     | 601     | 705    | 12    | 0          | 117    | 0      | 131    | 120    | 117    |
|       | E3 | 111       | 689     | 33      | 41      | 398    | 1     | 0          | 167    | 131    | 82     | 199    | 113    |
|       | L1 | 3709      | 13379   | 5309    | 5490    | 3717   | 116   | 11         | 314    | 131    | 556    | 1394   | 1389   |
| 7     | E1 | 6485      | 10992   | 4955    | 1363    | 1296   | 90    | 139        | 393    | 193    | 472    | 1092   | 3446   |
|       | E2 | 5181      | 11033   | 1362    | 1389    | 3405   | 0     | 364        | 180    | 624    | 738    | 1916   | 3056   |
|       | E3 | 9192      | 6329    | 511     | 447     | 746    | 0     | 379        | 310    | 4306   | 303    | 2225   | 1955   |
|       | L2 | 17779     | 18753   | 748     | 3899    | 5447   | 90    | 133        | 663    | 813    | 1113   | 3233   | 6887   |
| 8     | E1 | 145       | 323     | 341     | 59      | 200    | 0     | 0          | 168    | 145    | 93     | 171    | 149    |
|       | E2 | 3633      | 9668    | 1346    | 906     | 390    | 19    | 319        | 1460   | 1246   | 1499   | 1493   | 1962   |
|       | E3 | 18234     | 11925   | 1664    | 730     | 597    | 0     | 31         | 3111   | 1112   | 2492   | 1500   | 2937   |
|       | L3 | 19712     | 12926   | 2151    | 1646    | 1597   | 19    | 319        | 4739   | 4613   | 4074   | 4165   | 4997   |
| 9     | E1 | 13983     | 29613   | 31431   | 18838   | 4198   | 889   | 433        | 3673   | 6780   | 3302   | 10227  | 11938  |
|       | E2 | 1426      | 6607    | 1195    | 590     | 606    | 17    | 0          | 341    | 470    | 762    | 962    | 605    |
|       | E3 | 2994      | 2909    | 493     | 119     | 160    | 0     | 46         | 697    | 592    | 471    | 512    | 329    |
|       | L4 | 17343     | 36950   | 33019   | 18647   | 4964   | 716   | 479        | 6611   | 7912   | 10635  | 12101  | 12972  |
| 10    | E1 | 29619     | 611066  | 773697  | 393739  | 92467  | 9540  | 3393       | 202330 | 212046 | 238489 | 230004 | 215452 |
|       | E2 | 6484      | 13195   | 6457    | 3501    | 2203   | 191   | 14         | 2419   | 1776   | 3977   | 4267   | 4497   |
|       | E3 | 13917     | 19385   | 510     | 787     | 837    | 0     | 0          | 4854   | 4071   | 4071   | 3959   | 3207   |
|       | L5 | 113921    | 679386  | 777356  | 407996  | 96207  | 9731  | 3411       | 209603 | 219893 | 246437 | 239139 | 223176 |
| 11    | E1 | 6936      | 3311    | 7359    | 3939    | 9392   | 171   | 1697       | 12954  | 13967  | 16244  | 16309  | 17993  |
|       | E2 | 909       | 4950    | 1997    | 1962    | 1454   | 0     | 179        | 666    | 743    | 1115   | 1131   | 991    |
|       | E3 | 2492      | 3377    | 359     | 193     | 230    | 0     | 32         | 667    | 592    | 597    | 553    | 656    |
|       | L1 | 10344     | 41139   | 8714    | 41694   | 11966  | 171   | 1917       | 14187  | 15182  | 17956  | 17993  | 19630  |
| 12    | E1 | 1112      | 6749    | 1164    | 4542    | 2578   | 1     | 1362       | 130    | 703    | 656    | 1195   | 2285   |
|       | E2 | 5775      | 10839   | 9126    | 6986    | 11251  | 0     | 1613       | 646    | 2394   | 1704   | 1927   | 3938   |
|       | E3 | 9598      | 11971   | 1422    | 929     | 1290   | 0     | 436        | 1597   | 5608   | 1274   | 1598   | 3436   |
|       | L2 | 17467     | 19459   | 11612   | 12467   | 15109  | 10    | 3411       | 1373   | 8705   | 3634   | 4710   | 9659   |
| 13    | E1 | 419       | 1993    | 1751    | 1509    | 498    | 11    | 154        | 199    | 399    | 653    | 567    | 643    |
|       | E2 | 749       | 3724    | 1300    | 392     | 998    | 0     | 184        | 690    | 654    | 471    | 715    | 397    |
|       | E3 | 2023      | 3303    | 434     | 310     | 210    | 0     | 29         | 563    | 493    | 607    | 592    | 556    |
|       | L3 | 3333      | 5990    | 3485    | 1711    | 1706   | 11    | 166        | 1452   | 1535   | 1731   | 1974   | 3096   |
| 14    | E1 | 3592      | 17864   | 44186   | 21392   | 5090   | 162   | 148        | 6148   | 7136   | 9303   | 6910   | 10524  |
|       | E2 | 188       | 1438    | 1572    | 786     | 989    | 0     | 62         | 431    | 543    | 418    | 433    | 611    |
|       | E3 | 1455      | 1916    | 124     | 158     | 130    | 0     | 3          | 311    | 323    | 319    | 301    | 334    |
|       | L4 | 5633      | 19918   | 45982   | 21336   | 6099   | 162   | 153        | 6887   | 8102   | 10040  | 7644   | 11469  |
| 15    | E1 | 20871     | 469510  | 1017129 | 461993  | 103592 | 3164  | 17414      | 179307 | 192948 | 229002 | 229907 | 234517 |
|       | E2 | 11996     | 61900   | 32813   | 19777   | 13655  | 59    | 1023       | 9654   | 9709   | 11875  | 12992  | 14584  |
|       | E3 | 30293     | 36166   | 3700    | 1179    | 2125   | 0     | 156        | 7520   | 7336   | 7349   | 6795   | 7035   |
|       | L5 | 132970    | 567576  | 1353442 | 462949  | 119372 | 3323  | 18593      | 195481 | 209893 | 247226 | 249694 | 256136 |
| TOTAL |    | 622599    | 1338416 | 2384535 | 1131952 | 369586 | 18719 | 47999      | 598313 | 636423 | 709153 | 697055 | 699735 |

Fonte: Censos Demográficos de 1970, 1980 e 1991.

continua ...

TABELA 5: continuação

|       |    | RENDIA |        |        |        |        |        | CORIA   |        | TOTAL   |
|-------|----|--------|--------|--------|--------|--------|--------|---------|--------|---------|
|       |    | R5     | R7     | R7     | R9     | R10    | R11    | E1      | E2     |         |
| 11    | E1 | 343    | 616    | 334    | 1639   | 1271   | 829    | 1133    | 789    | 33874   |
|       | E2 | 1      | 332    | 1      | 1      | 574    | 117    | 1118    | 303    | 4184    |
|       | E3 | 1      | 121    | 111    | 111    | 111    | 111    | 819     | 1      | 1917    |
| E1    |    | 343    | 1119   | 1178   | 1738   | 1945   | 844    | 2954    | 1161   | 10745   |
| 12    | E1 | 775    | 3713   | 3299   | 1399   | 3485   | 199    | 11789   | 3819   | 48931   |
|       | E2 | 118    | 1303   | 3199   | 1311   | 4196   | 801    | 11136   | 3317   | 19036   |
|       | E3 | 1199   | 1392   | 1114   | 1115   | 789    | 491    | 7113    | 4117   | 33660   |
| E1    |    | 1711   | 3399   | 7491   | 6824   | 1479   | 1810   | 41347   | 11962  | 117517  |
| 13    | E1 | 1      | 302    | 119    | 199    | 1      | 0      | 1790    | 303    | 8009    |
|       | E2 | 1113   | 1110   | 779    | 1414   | 510    | 496    | 3409    | 1613   | 10786   |
|       | E3 | 1099   | 1116   | 1515   | 1199   | 1710   | 1304   | 17118   | 3894   | 81327   |
| E1    |    | 1102   | 3016   | 3474   | 3713   | 1110   | 790    | 29324   | 3400   | 101172  |
| 14    | E1 | 4774   | 6819   | 7349   | 8479   | 3919   | 8343   | 43396   | 7801   | 151191  |
|       | E2 | 114    | 196    | 790    | 124    | 308    | 834    | 3968    | 694    | 13956   |
|       | E3 | 36     | 110    | 773    | 832    | 309    | 694    | 3963    | 1196   | 11777   |
| E1    |    | 8702   | 1866   | 1910   | 7394   | 4118   | 6581   | 71117   | 2491   | 179924  |
| 15    | E1 | 139681 | 111696 | 110184 | 114630 | 24583  | 116978 | 1106379 | 301204 | 4119722 |
|       | E2 | 1491   | 164    | 1704   | 1137   | 1710   | 8360   | 18714   | 1719   | 61916   |
|       | E3 | 1097   | 3715   | 1510   | 1713   | 1876   | 3536   | 27968   | 7771   | 117208  |
| E1    |    | 143993 | 114696 | 118489 | 110817 | 29949  | 113974 | 1133049 | 311694 | 4399946 |
| 16    | E1 | 1043   | 3818   | 6685   | 6864   | 871    | 139    | 29419   | 1305   | 91100   |
|       | E2 | 124    | 399    | 407    | 134    | 1164   | 94     | 3679    | 177    | 11969   |
|       | E3 | 86     | 87     | 183    | 156    | 483    | 0      | 1486    | 135    | 8160    |
| E1    |    | 1031   | 4050   | 7315   | 7998   | 7485   | 131    | 34593   | 1417   | 100130  |
| 17    | E1 | 1900   | 4449   | 3799   | 3898   | 1951   | 93     | 19903   | 3879   | 67143   |
|       | E2 | 1711   | 3019   | 1967   | 3951   | 781    | 173    | 10778   | 3136   | 71983   |
|       | E3 | 1120   | 1416   | 1131   | 1309   | 1501   | 140    | 14763   | 3761   | 32572   |
| E1    |    | 6042   | 8993   | 8176   | 8184   | 13073  | 506    | 55341   | 8535   | 101896  |
| 18    | E1 | 32     | 47     | 114    | 118    | 94     | 0      | 1159    | 110    | 3537    |
|       | E2 | 1662   | 1706   | 1766   | 1179   | 1502   | 169    | 15268   | 1613   | 50643   |
|       | E3 | 1969   | 2631   | 2983   | 3032   | 2632   | 361    | 24250   | 3940   | 84170   |
| E1    |    | 4723   | 4404   | 4743   | 4956   | 4169   | 330    | 41577   | 3573   | 138480  |
| 19    | E1 | 15831  | 20190  | 20497  | 16422  | 7393   | 1998   | 113486  | 8139   | 379872  |
|       | E2 | 1115   | 1098   | 1153   | 1618   | 1514   | 123    | 9031    | 510    | 18653   |
|       | E3 | 843    | 616    | 303    | 649    | 1261   | 35     | 5662    | 992    | 13932   |
| E1    |    | 17379  | 21104  | 21933  | 18899  | 10268  | 1136   | 133178  | 9641   | 428457  |
| 20    | E1 | 117139 | 198121 | 177532 | 157599 | 102910 | 20053  | 1782342 | 232729 | 5988201 |
|       | E2 | 4568   | 5209   | 4688   | 4799   | 3996   | 462    | 39919   | 2493   | 134203  |
|       | E3 | 5076   | 3290   | 3682   | 3195   | 3418   | 255    | 31404   | 4684   | 111164  |
| E1    |    | 124993 | 206699 | 185872 | 165652 | 139123 | 20770  | 1933664 | 239990 | 6233668 |
| 21    | E1 | 17417  | 18972  | 19033  | 16399  | 14241  | 6362   | 183093  | 11866  | 404875  |
|       | E2 | 1112   | 1179   | 1424   | 1374   | 1831   | 156    | 11713   | 469    | 36544   |
|       | E3 | 839    | 675    | 612    | 731    | 732    | 247    | 6174    | 425    | 19798   |
| E1    |    | 19049  | 19416  | 16976  | 18554  | 16804  | 7065   | 179390  | 12760  | 301217  |
| 22    | E1 | 1091   | 3375   | 3799   | 4371   | 3743   | 356    | 21757   | 3299   | 73167   |
|       | E2 | 1347   | 4504   | 6363   | 7513   | 19484  | 944    | 49042   | 3935   | 158631  |
|       | E3 | 1931   | 1361   | 1082   | 3022   | 3303   | 154    | 20297   | 2795   | 75046   |
| E1    |    | 6499   | 9940   | 11194  | 13906  | 16650  | 1554   | 83096   | 9929   | 316044  |
| 23    | E1 | 609    | 928    | 1184   | 865    | 486    | 174    | 6729    | 172    | 11003   |
|       | E2 | 858    | 772    | 829    | 811    | 1162   | 198    | 7614    | 399    | 14006   |
|       | E3 | 912    | 548    | 752    | 733    | 729    | 177    | 5861    | 429    | 18840   |
| E1    |    | 1660   | 1248   | 1726   | 2599   | 2376   | 349    | 20204   | 1090   | 63849   |
| 24    | E1 | 8316   | 9414   | 10291  | 9315   | 7083   | 3780   | 86331   | 5844   | 276526  |
|       | E2 | 634    | 665    | 609    | 345    | 785    | 143    | 5994    | 176    | 19910   |
|       | E3 | 260    | 430    | 342    | 440    | 451    | 179    | 3424    | 259    | 11049   |
| E1    |    | 10210  | 11509  | 11293  | 10610  | 8319   | 4202   | 95749   | 6379   | 306385  |
| 25    | E1 | 219849 | 221990 | 206399 | 197612 | 143175 | 87540  | 1993937 | 162222 | 6439478 |
|       | E2 | 13597  | 14192  | 15189  | 15603  | 14624  | 5679   | 132451  | 6259   | 41613   |
|       | E3 | 8599   | 6912   | 7504   | 7573   | 7461   | 2223   | 68178   | 6285   | 12338   |
| E1    |    | 139735 | 142994 | 129091 | 110789 | 165260 | 95442  | 1194566 | 174766 | 7077997 |
| TOTAL |    | 699403 | 675158 | 645599 | 600000 | 489127 | 268265 | 5932339 | 808460 | 2022019 |

Fonte: Censos Demográficos de 1970, 1980 e 1991.

Associadas à tabela 5 temos, então, a matriz de dados,  $N_1$ , e a matriz dos centróides dos grupos,  $N_2$ . A matriz de dados, apresentada a seguir, é a matriz cujos perfis linha foram tomados como pontos suplementares representando os elementos dos grupos, onde os eixos são determinados pelos centróides destes grupos. Os elementos desta matriz nos dão, para cada localização doméstica, o número de indivíduos em cada etapa de vida e sua distribuição para cada variável definida nas colunas.

$N_1 =$

|       |        |         |        |        |      |     |   |       |        |        |        |        |        |        |        |        |        |        |
|-------|--------|---------|--------|--------|------|-----|---|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 50    | 1822   | 97      | 24     | 3423   | 1362 | 117 | 1 | 421   | 0      | 0      | 943    | 613    | 954    | 1639   | 1071   | 527    | 7100   | 759    |
| 503   | 897    | 0       | 0      | 117    | 0    | 0   | 0 | 10    | 0      | 0      | 20     | 135    | 0      | 0      | 0      | 0      | 0      | 0      |
| 121   | 415    | 0       | 0      | 110    | 0    | 0   | 0 | 0     | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 0      |
| 5200  | 6134   | 1239    | 1135   | 1489   | 532  | 0   | 0 | 509   | 515    | 398    | 435    | 737    | 3703   | 3234   | 2999   | 2485   | 199    | 11759  |
| 5398  | 7209   | 438     | 1138   | 1300   | 198  | 0   | 0 | 513   | 509    | 199    | 1339   | 515    | 1303   | 3199   | 3321   | 4195   | 911    | 12135  |
| 6614  | 5306   | 110     | 11     | 239    | 1009 | 102 | 0 | 438   | 432    | 535    | 1213   | 1499   | 1302   | 1134   | 1005   | 799    | 461    | 7103   |
| 401   | 1235   | 100     | 0      | 297    | 0    | 0   | 0 | 403   | 405    | 39     | 199    | 297    | 0      | 0      | 0      | 0      | 0      | 1700   |
| 2900  | 8701   | 509     | 901    | 710    | 99   | 0   | 0 | 513   | 1113   | 1274   | 1511   | 509    | 1003   | 1231   | 799    | 1404   | 500    | 895    |
| 3601  | 10100  | 177     | 510    | 421    | 110  | 0   | 0 | 297   | 895    | 1137   | 1404   | 1513   | 4212   | 1999   | 1206   | 1515   | 1496   | 1804   |
| 3142  | 29699  | 3153    | 3944   | 1859   | 411  | 0   | 0 | 2968  | 2908   | 3084   | 4229   | 4121   | 4702   | 9829   | 7549   | 6479   | 2919   | 5343   |
| 1038  | 1718   | 501     | 397    | 119    | 0    | 0   | 0 | 94    | 397    | 411    | 431    | 334    | 104    | 106    | 74     | 294    | 808    | 534    |
| 1354  | 1355   | 105     | 110    | 135    | 0    | 0   | 0 | 79    | 94     | 371    | 421    | 112    | 112    | 76     | 142    | 773    | 552    | 694    |
| 12765 | 7477   | 119427  | 154135 | 210161 | 631  | 0   | 0 | 13587 | 135499 | 140890 | 151355 | 138544 | 138336 | 137552 | 117606 | 117134 | 114680 | 84553  |
| 5520  | 12769  | 59      | 1036   | 951    | 110  | 0   | 0 | 159   | 753    | 1519   | 1929   | 1271   | 1015   | 1491   | 164    | 2794   | 1137   | 1701   |
| 10311 | 12381  | 135     | 335    | 1222   | 0    | 0   | 0 | 284   | 4949   | 4815   | 3871   | 3189   | 1065   | 1350   | 3115   | 1500   | 1713   | 2575   |
| 3962  | 11431  | 6355    | 5349   | 1710   | 64   | 0   | 0 | 150   | 31     | 343    | 865    | 1144   | 3343   | 5605   | 6565   | 6964   | 5571   | 139    |
| 471   | 1201   | 411     | 611    | 715    | 52   | 0   | 0 | 27    | 132    | 329    | 117    | 194    | 368    | 437    | 834    | 1454   | 94     | 3679   |
| 480   | 659    | 103     | 41     | 308    | 0    | 0   | 0 | 167   | 151    | 91     | 199    | 129    | 63     | 57     | 153    | 259    | 460    | 0      |
| 3435  | 10921  | 4055    | 1163   | 1196   | 30   | 0   | 0 | 190   | 213    | 193    | 471    | 1692   | 2446   | 1910   | 4448   | 3739   | 3515   | 2951   |
| 5162  | 11323  | 1582    | 1389   | 8415   | 0    | 0   | 0 | 364   | 150    | 694    | 739    | 1916   | 3105   | 1702   | 2119   | 2957   | 3351   | 7621   |
| 3192  | 6329   | 511     | 447    | 745    | 0    | 0   | 0 | 379   | 310    | 4326   | 903    | 2225   | 1055   | 1425   | 1425   | 1431   | 1308   | 2501   |
| 145   | 513    | 342     | 50     | 121    | 0    | 0   | 0 | 169   | 145    | 83     | 172    | 148    | 32     | 47     | 104    | 115    | 84     | 0      |
| 3353  | 9639   | 1946    | 416    | 891    | 29   | 0   | 0 | 319   | 1459   | 1245   | 1499   | 1493   | 1962   | 1632   | 1705   | 1765   | 1178   | 1502   |
| 13234 | 11305  | 1654    | 730    | 597    | 0    | 0   | 0 | 30    | 3111   | 2222   | 2492   | 3500   | 2977   | 1969   | 1631   | 2853   | 3362   | 3622   |
| 12963 | 55515  | 34431   | 19239  | 4139   | 629  | 0   | 0 | 433   | 5573   | 6750   | 8302   | 10227  | 11838  | 15331  | 31190  | 24487  | 16422  | 7693   |
| 1416  | 5307   | 1135    | 500    | 616    | 17   | 0   | 0 | 341   | 411    | 762    | 952    | 515    | 1105   | 1099   | 1063   | 1629   | 1324   | 123    |
| 2954  | 1329   | 483     | 219    | 160    | 0    | 0   | 0 | 45    | 697    | 592    | 471    | 812    | 529    | 843    | 826    | 393    | 549    | 1261   |
| 88619 | 641006 | 779697  | 363739 | 20487  | 8541 | 0   | 0 | 3365  | 202330 | 111045 | 138489 | 230004 | 215451 | 217039 | 199121 | 177830 | 157535 | 122310 |
| 5454  | 33195  | 6657    | 3501   | 2213   | 151  | 0   | 0 | 2419  | 3776   | 3877   | 4267   | 4497   | 4555   | 5209   | 4659   | 4750   | 3695   | 462    |
| 12907 | 15195  | 302     | 757    | 537    | 0    | 0   | 0 | 4854  | 4071   | 4071   | 3559   | 3227   | 3075   | 3290   | 3682   | 3295   | 3418   | 355    |
| 6996  | 33812  | 73759   | 39339  | 9892   | 271  | 0   | 0 | 1627  | 12854  | 13867  | 16144  | 16309  | 17893  | 17427  | 16572  | 15133  | 16389  | 14241  |
| 909   | 4950   | 1007    | 1062   | 1454   | 0    | 0   | 0 | 179   | 665    | 713    | 1115   | 1131   | 91     | 1130   | 1179   | 1124   | 1374   | 1531   |
| 1439  | 3377   | 359     | 193    | 130    | 0    | 0   | 0 | 51    | 657    | 520    | 377    | 553    | 655    | 530    | 675    | 519    | 791    | 730    |
| 2012  | 6749   | 1164    | 4542   | 2075   | 17   | 0   | 0 | 1362  | 130    | 713    | 656    | 1195   | 2235   | 2191   | 3375   | 3149   | 4371   | 3743   |
| 5775  | 20639  | 9216    | 6946   | 11251  | 0    | 0   | 0 | 1613  | 646    | 1394   | 1704   | 1317   | 3959   | 1947   | 4514   | 5253   | 7513   | 19484  |
| 3580  | 11371  | 1422    | 109    | 1291   | 0    | 0   | 0 | 435   | 1537   | 5608   | 1274   | 1528   | 3455   | 1851   | 1951   | 2132   | 2122   | 3323   |
| 419   | 1913   | 1721    | 1509   | 498    | 11   | 0   | 0 | 154   | 105    | 398    | 653    | 557    | 643    | 500    | 929    | 1154   | 855    | 486    |
| 388   | 3724   | 1510    | 422    | 909    | 0    | 0   | 0 | 154   | 690    | 654    | 471    | 715    | 997    | 658    | 772    | 820    | 911    | 1162   |
| 1713  | 3393   | 154     | 310    | 111    | 0    | 0   | 0 | 563   | 493    | 607    | 522    | 555    | 512    | 548    | 752    | 713    | 729    | 177    |
| 3592  | 17554  | 44296   | 21392  | 5080   | 162  | 0   | 0 | 848   | 6145   | 7235   | 9303   | 8910   | 17524  | 9326   | 9414   | 10091  | 9325   | 7083   |
| 925   | 2438   | 1872    | 745    | 899    | 0    | 0   | 0 | 62    | 431    | 543    | 418    | 433    | 611    | 634    | 665    | 600    | 845    | 785    |
| 1455  | 1916   | 124     | 159    | 130    | 0    | 0   | 0 | 3     | 311    | 323    | 319    | 301    | 334    | 350    | 430    | 342    | 440    | 451    |
| 30571 | 483510 | 1017229 | 461955 | 103692 | 3254 | 0   | 0 | 17414 | 178307 | 180948 | 229022 | 219007 | 234817 | 116549 | 221390 | 206359 | 187612 | 143175 |
| 11806 | 61900  | 30513   | 19777  | 13655  | 59   | 0   | 0 | 1023  | 9654   | 9709   | 11875  | 12992  | 14584  | 13587  | 14192  | 15188  | 15303  | 14624  |
| 30293 | 36165  | 3700    | 2179   | 1111   | 0    | 0   | 0 | 156   | 7520   | 7535   | 7349   | 6795   | 7035   | 6599   | 6912   | 7504   | 7573   | 7461   |

A matriz dos centróides dos grupos,  $N_2$ , como enfatizado no capítulo anterior, é a matriz cujos perfis linha são os representantes dos grupos na nuvem e são usados para determinar as direções dos eixos. Desta forma, foi aplicado o algoritmo da AC, como descrito no capítulo 2, na matriz  $N_2$ . Esta matriz é apresentada a seguir e cada um de seus elementos nos dá o número de indivíduos em cada localização doméstica, desconsiderando as etapas de vida, distribuídos segundo cada variável definida nas colunas.

$N_2 =$

|        |        |         |        |        |       |       |        |        |        |        |        |        |        |        |        |        |        |         |        |
|--------|--------|---------|--------|--------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|--------|
| 1376   | 1387   | 147     | 324    | 373    | 1362  | 117   | 0      | 520    | 158    | 117    | 343    | 1119   | 1078   | 1738   | 1945   | 644    | 854    | 1762    |        |
| 17412  | 17308  | 1387    | 3214   | 3588   | 13399 | 131   | 1810   | 1115   | 1193   | 117    | 1711   | 7395   | 7481   | 6214   | 7479   | 1500   | 3157   | 11431   |        |
| 12894  | 17426  | 835     | 1131   | 1328   | 301   | 326   | 1901   | 3815   | 1704   | 3213   | 1136   | 3102   | 3610   | 3404   | 3700   | 2202   | 1700   | 13204   | 5439   |
| 11134  | 34132  | 8398    | 4131   | 3111   | 411   | 303   | 1751   | 3576   | 3806   | 4892   | 4867   | 5012   | 7156   | 8712   | 7334   | 4113   | 6561   | 5117    | 8421   |
| 133196 | 773117 | 303640  | 165357 | 84383  | 743   | 5779  | 141131 | 147193 | 156616 | 143304 | 141197 | 143095 | 114555 | 119459 | 119549 | 88849  | 113674 | 1152049 | 311694 |
| 3139   | 13379  | 7484    | 6490   | 3717   | 116   | 13    | 344    | 181    | 556    | 1234   | 1389   | 3512   | 6153   | 7115   | 7356   | 7485   | 12     | 34533   | 1317   |
| 17719  | 28753  | 7449    | 3899   | 5447   | 0     | 1033  | 633    | 5113   | 1113   | 8233   | 9587   | 6141   | 9193   | 9176   | 8194   | 13073  | 516    | 55341   | 5215   |
| 17343  | 11946  | 3851    | 1346   | 1837   | 08    | 348   | 4739   | 4613   | 4074   | 4165   | 4987   | 4713   | 4404   | 4743   | 4356   | 4165   | 530    | 40577   | 5573   |
| 17343  | 68850  | 36209   | 16647  | 4954   | 716   | 479   | 6311   | 7812   | 10835  | 12001  | 11972  | 17379  | 22104  | 11333  | 15529  | 10168  | 1136   | 133179  | 3641   |
| 113980 | 673196 | 773356  | 367996 | 85217  | 8731  | 3411  | 109616 | 118893 | 246437 | 136150 | 223176 | 114890 | 206309 | 186872 | 165952 | 139123 | 20770  | 1833664 | 239922 |
| 11344  | 41139  | 78214   | 41394  | 11166  | 171   | 1317  | 14137  | 15382  | 17955  | 17993  | 19530  | 12049  | 18416  | 16376  | 18554  | 16994  | 7065   | 170990  | 12763  |
| 17367  | 39459  | 17611   | 11457  | 15119  | 17    | 3411  | 1373   | 3715   | 3354   | 4710   | 3659   | 6459   | 9940   | 11134  | 13806  | 23550  | 1554   | 83066   | 3929   |
| 3530   | 3630   | 4635    | 1711   | 1716   | 11    | 436   | 1451   | 1535   | 1731   | 1874   | 2196   | 1651   | 2143   | 2716   | 1599   | 1376   | 549    | 2104    | 1980   |
| 5632   | 11313  | 45962   | 22336  | 6199   | 162   | 313   | 8887   | 8112   | 10040  | 8644   | 11482  | 10100  | 10509  | 11133  | 10610  | 8319   | 4202   | 95749   | 6379   |
| 121670 | 567576 | 1153442 | 403948 | 118371 | 3323  | 15563 | 156481 | 109893 | 247226 | 249694 | 255136 | 138735 | 242994 | 229091 | 210736 | 185260 | 95442  | 2184565 | 174766 |

Na próxima seção, descrevemos como, a partir destes dados, obtemos as coordenadas dos pontos a serem analisados.

### 4.3 TRANSFORMAÇÃO DOS DADOS

Na seção anterior apresentamos a tabela 5, que é a tabela de dados à qual associamos as matrizes  $N$  e  $N_2$ . A matriz  $N$  é a matriz em cujas categorias não foi feito qualquer agrupamento. A matriz  $N_2$  é a matriz obtida a partir do agrupamento das linhas. Como o nosso interesse aqui é o de analisar a variável Localização Doméstica, como descrita em cada categoria, pelas variáveis descritoras, o algoritmo da AC foi aplicado na matriz  $N_2$  e os perfis linha de  $N$ , colocados como pontos suplementares. Nesta seção mostramos, a partir dessas matrizes, a transformação desses dados até a obtenção das coordenadas dos pontos.

#### 4.3.1 APLICAÇÃO DO ALGORITMO DA AC

Inicialmente, transformamos as matrizes de freqüências absolutas em matrizes de perfis. Deste modo, de  $N$  obtemos os perfis das etapas de vida ao longo das variáveis descritoras, que serão tratadas como pontos suplementares. De  $N_2$  obtemos, através da expressão (2.3.3), os perfis das localizações domésticas ao longo das variáveis descritoras, que são as linhas de  $R_2$ , dada a seguir.

$$R_2 = \begin{bmatrix} 0.0462 & 0.1267 & 0.0318 & 0.0310 & 0.1254 & 0.0000 & 0.0457 & 0.0039 & 0.0000 & 0.0174 & 0.0079 & 0.0072 & 0.0317 & 0.0376 & 0.0361 & 0.0584 & 0.0653 & 0.0216 & 0.2976 & 0.0356 \\ 0.1365 & 0.1357 & 0.0147 & 0.0181 & 0.0281 & 0.0000 & 0.0203 & 0.0098 & 0.0118 & 0.0066 & 0.0093 & 0.1243 & 0.0212 & 0.0580 & 0.0587 & 0.0495 & 0.0556 & 0.0117 & 0.2434 & 0.0898 \\ 0.1265 & 0.1702 & 0.0488 & 0.0198 & 0.0139 & 0.0000 & 0.0039 & 0.0178 & 0.0377 & 0.0267 & 0.0016 & 0.0014 & 0.0296 & 0.0356 & 0.0336 & 0.0365 & 0.0217 & 0.0266 & 0.2729 & 0.0933 \\ 0.0622 & 0.1213 & 0.0287 & 0.0231 & 0.0173 & 0.0022 & 0.0017 & 0.0194 & 0.0205 & 0.0218 & 0.0273 & 0.0277 & 0.0279 & 0.0400 & 0.0430 & 0.0410 & 0.0230 & 0.0266 & 0.2803 & 0.0530 \\ 0.0317 & 0.1781 & 0.0685 & 0.0377 & 0.0192 & 0.0001 & 0.0013 & 0.0321 & 0.0335 & 0.0356 & 0.0325 & 0.0321 & 0.0308 & 0.0293 & 0.0285 & 0.0272 & 0.0202 & 0.0282 & 0.2624 & 0.0710 \\ 0.0473 & 0.1224 & 0.0283 & 0.0584 & 0.0340 & 0.0010 & 0.0001 & 0.0031 & 0.0016 & 0.0050 & 0.0117 & 0.0127 & 0.0329 & 0.0553 & 0.0689 & 0.0726 & 0.0685 & 0.0021 & 0.3166 & 0.0166 \\ 0.0325 & 0.1500 & 0.0414 & 0.0203 & 0.0284 & 0.0004 & 0.0053 & 0.0034 & 0.0272 & 0.0110 & 0.0273 & 0.0290 & 0.0315 & 0.0422 & 0.0426 & 0.0426 & 0.0682 & 0.0026 & 0.2888 & 0.0444 \\ 0.1230 & 0.1582 & 0.0079 & 0.0118 & 0.0115 & 0.0002 & 0.0025 & 0.0342 & 0.0333 & 0.0234 & 0.0300 & 0.0360 & 0.0341 & 0.0318 & 0.0342 & 0.0329 & 0.0308 & 0.0038 & 0.2930 & 0.0402 \\ 0.0404 & 0.1562 & 0.0945 & 0.0388 & 0.0115 & 0.0016 & 0.0011 & 0.0154 & 0.0182 & 0.0245 & 0.0280 & 0.0302 & 0.0405 & 0.0515 & 0.0511 & 0.0433 & 0.0239 & 0.0049 & 0.3108 & 0.0225 \\ 0.0183 & 0.1091 & 0.1065 & 0.0623 & 0.0153 & 0.0015 & 0.0005 & 0.0336 & 0.0351 & 0.0326 & 0.0382 & 0.0358 & 0.0361 & 0.0332 & 0.0298 & 0.0266 & 0.0209 & 0.0033 & 0.2947 & 0.0385 \\ 0.0187 & 0.0784 & 0.1437 & 0.0738 & 0.0200 & 0.0004 & 0.0034 & 0.0257 & 0.0277 & 0.0325 & 0.0326 & 0.0354 & 0.0345 & 0.0334 & 0.0307 & 0.0336 & 0.0304 & 0.0128 & 0.3101 & 0.0231 \\ 0.0567 & 0.1289 & 0.0575 & 0.0407 & 0.0493 & 0.0000 & 0.0111 & 0.0077 & 0.0284 & 0.0118 & 0.0153 & 0.0315 & 0.0212 & 0.0321 & 0.0365 & 0.0454 & 0.0867 & 0.0090 & 0.3041 & 0.0291 \\ 0.0521 & 0.1382 & 0.0733 & 0.0424 & 0.0267 & 0.0003 & 0.0068 & 0.0227 & 0.0240 & 0.0271 & 0.0293 & 0.0128 & 0.0259 & 0.0352 & 0.0426 & 0.0407 & 0.0372 & 0.0085 & 0.3164 & 0.0169 \\ 0.0183 & 0.0715 & 0.1500 & 0.0729 & 0.0139 & 0.0005 & 0.0029 & 0.0224 & 0.0264 & 0.0327 & 0.0314 & 0.0274 & 0.0332 & 0.0343 & 0.0366 & 0.0346 & 0.0271 & 0.0137 & 0.3125 & 0.0208 \\ 0.0187 & 0.0601 & 0.1488 & 0.0682 & 0.0168 & 0.0004 & 0.0026 & 0.0276 & 0.0296 & 0.0349 & 0.0352 & 0.0361 & 0.0337 & 0.0343 & 0.0323 & 0.0297 & 0.0233 & 0.0134 & 0.3086 & 0.0246 \end{bmatrix}$$

Esta matriz nos mostra, por linhas, 15 (quinze) pontos do espaço  $R^{20}$ , que são as 5 (cinco) localizações domésticas nos 3 (três) censos. Também de  $N_2$ , obtemos os perfis das variáveis descritoras ao longo destas localizações domésticas, que são as colunas da matriz  $C_2'$ , apresentada a seguir. A notação ( $'$ ) em  $C_2'$  indica que a matriz abaixo está transposta, com relação à apresentada no desenvolvimento teórico.

$$C_2' = \begin{bmatrix} 0.0026 & 0.0042 & 1.0004 & 0.0008 & 0.1113 & 1.0000 & 0.1022 & 0.1001 & 0.1000 & 0.1007 & 0.0008 & 1.1002 & 0.0013 & 0.0016 & 0.1008 & 0.0008 & 0.0009 & 0.0014 & 0.0114 & 0.0013 \\ 0.0033 & 0.0074 & 0.0008 & 0.0020 & 0.1009 & 0.0000 & 0.0000 & 0.0001 & 0.1003 & 0.0015 & 0.0017 & 1.1004 & 0.0009 & 0.0010 & 0.1008 & 0.0005 & 0.0005 & 0.0052 & 0.0052 & 0.0141 \\ 0.0045 & 0.0074 & 0.0003 & 0.0009 & 0.1003 & 0.0003 & 0.0006 & 0.0000 & 0.0009 & 0.0008 & 0.0045 & 1.1005 & 0.0043 & 0.0053 & 0.0052 & 0.1001 & 0.0045 & 0.0100 & 0.0047 & 0.0066 \\ 0.0013 & 0.0147 & 0.0017 & 0.0005 & 0.1009 & 0.0020 & 0.0074 & 0.0046 & 0.0057 & 0.0055 & 0.0070 & 1.1007 & 0.0072 & 0.0106 & 0.0103 & 0.0122 & 0.0084 & 0.0244 & 0.0064 & 0.0117 \\ 0.0063 & 0.0011 & 0.0276 & 0.1438 & 0.1045 & 0.0470 & 0.1409 & 0.0198 & 0.0211 & 0.0208 & 0.0050 & 1.1009 & 0.0091 & 0.0145 & 0.1043 & 0.1000 & 0.1016 & 0.4617 & 0.1041 & 0.0055 \\ 0.0100 & 0.0017 & 0.0021 & 0.1006 & 0.1010 & 0.0070 & 0.0003 & 0.0005 & 0.0002 & 0.0007 & 0.0015 & 1.1010 & 0.0052 & 0.0089 & 0.0113 & 0.0102 & 0.0153 & 0.0008 & 0.0058 & 0.0020 \\ 0.0039 & 0.0113 & 0.0039 & 0.0033 & 0.1001 & 0.0057 & 0.0051 & 0.0011 & 0.0081 & 0.0009 & 0.0070 & 1.1011 & 0.0087 & 0.0119 & 0.0105 & 0.0106 & 0.0267 & 0.0018 & 0.0093 & 0.0105 \\ 0.0025 & 0.0094 & 0.0016 & 0.0014 & 0.0044 & 0.0017 & 0.0084 & 0.0080 & 0.0070 & 0.0057 & 0.0059 & 1.1012 & 0.0068 & 0.0065 & 0.0073 & 0.0075 & 0.0087 & 0.0019 & 0.0068 & 0.0068 \\ 0.0031 & 0.0038 & 0.0153 & 0.0114 & 0.1017 & 0.0455 & 0.0116 & 0.0112 & 0.0122 & 0.0148 & 0.0172 & 1.1013 & 0.0052 & 0.0027 & 0.0039 & 0.0009 & 0.0209 & 0.0079 & 0.0024 & 0.0119 \\ 0.0101 & 0.0008 & 0.0044 & 0.0046 & 0.0047 & 0.0190 & 0.0031 & 0.0032 & 0.0439 & 0.0475 & 0.0415 & 1.1014 & 0.0066 & 0.0060 & 0.0079 & 0.0060 & 0.0060 & 0.0074 & 0.0090 & 0.0067 \\ 0.0197 & 0.0188 & 0.0006 & 0.0053 & 0.1006 & 0.0172 & 0.0467 & 0.0241 & 0.0240 & 0.0253 & 0.0258 & 1.1015 & 0.0076 & 0.0072 & 0.0082 & 0.0009 & 0.0043 & 0.0261 & 0.0088 & 0.0157 \\ 0.0032 & 0.0168 & 0.0074 & 0.0108 & 0.1019 & 0.0006 & 0.0031 & 0.0040 & 0.0106 & 0.0051 & 0.0087 & 1.1016 & 0.0094 & 0.0145 & 0.0173 & 0.0031 & 0.0542 & 0.0057 & 0.0118 & 0.0110 \\ 0.0043 & 0.0017 & 0.0019 & 0.0023 & 0.1047 & 0.0013 & 0.0106 & 0.0024 & 0.0024 & 0.0024 & 0.0003 & 1.1017 & 0.0024 & 0.0033 & 0.0042 & 0.0043 & 0.0048 & 0.0020 & 0.0034 & 0.0013 \\ 0.0107 & 0.0093 & 0.0105 & 0.0183 & 0.1019 & 0.0132 & 0.0020 & 0.0117 & 0.0127 & 0.0141 & 0.0138 & 1.1018 & 0.0148 & 0.0185 & 0.0174 & 0.0176 & 0.0170 & 0.0150 & 0.0161 & 0.0078 \\ 0.0038 & 0.0433 & 0.4474 & 0.4192 & 0.0011 & 0.1114 & 0.4534 & 0.0011 & 0.0008 & 0.0486 & 0.0081 & 1.1019 & 0.0467 & 0.0099 & 0.0048 & 0.0013 & 0.0078 & 0.0057 & 0.0082 & 0.0131 \end{bmatrix}$$

Esta matriz nos mostra, por colunas, 20 (vinte) pontos do espaço  $R^{15}$ , sendo que os 6 (seis) primeiros representam as categorias de instrução, os 12 (doze) seguintes representam as categorias de renda e os 2 (dois) últimos, as localizações do domicílio.

A partir dessas matrizes, centralizando, ponderando e transformando os perfis, obtemos a matriz  $X_2$ , como dada em (2.3.22). Como comentamos no capítulo 2, podemos trabalhar apenas com os perfis linha, pois assim obtemos as coordenadas dos perfis linha e também as coordenadas dos perfis coluna, no subespaço ajustado. Isso acontece quando fazemos a decomposição em valores singulares da matriz  $X_2$ , como dado em (2.5.11). Esta decomposição nos dá, na matriz  $V$ , uma base do subespaço ajustado dos perfis linha, e em  $U$ , uma base do subespaço ajustado dos perfis coluna. A partir destas bases obtemos, através da expressão (2.3.33), as coordenadas principais das localizações domésticas dadas em  $F_2$  a seguir.

$$F_2 = \begin{bmatrix} 0.3544955 & 0.2242045 & -0.104671 & -0.432703 & 0.2667015 & 0.4622301 & 0.3389921 & 0.0749222 & -0.001175 & -0.06303 & -0.00954 & 0.0059368 & 0.0024552 & -0.000944 & 0.0763117 & 0.0583659 & -0.12365 & 0.1149594 & 0.2222701 & 0.117333 & -0.016333 & -0.016348 & -0.004408 & 0.0250556 & 0.0008414 & -0.000995 & -0.001254 & -0.000114 & 0.0104862 & 0.2253304 & -0.106721 & 0.0296664 & 0.1243713 & -0.049997 & 0.009714 & 0.0413644 & 0.0515282 & -0.015665 & 0.0361997 & -0.002246 & -0.002733 & -0.006106 & 0.0540182 & 0.0723604 & 0.0594465 & 0.1141158 & -0.143963 & 0.0331935 & 0.0321612 & 0.0140714 & 0.0782367 & 0.022562 & -0.017676 & 0.0131503 & -0.001866 & 0.0009372 & 0.0295126 & -0.131912 & 0.0452888 & -0.003991 & -0.004142 & -0.004983 & -0.003228 & -0.000814 & -0.003847 & -0.000772 & 0.0003285 & -0.000022 & 0.0001645 & 0.000069 & 0.0125651 & 0.4686141 & -0.040315 & -0.108489 & -0.034709 & -0.006818 & -0.00578 & 0.0282178 & -0.026317 & -0.000836 & -0.000136 & -0.008234 & -0.0004014 & -0.002751 & 0.4009893 & 0.4584301 & -0.105817 & -0.102211 & -0.000888 & -0.0090382 & -0.003203 & -0.008107 & 0.044399 & -0.057654 & -0.008496 & 0.003842 & -0.000012 & 0.0009311 & -0.025719 & 0.2089886 & -0.021492 & 0.1040771 & 0.050435 & -0.000084 & 0.0101849 & 0.0087423 & -0.0069747 & 0.0139222 & -0.004934 & 0.0140592 & 0.0015179 & -0.002725 & -0.107636 & 0.103215 & -0.024709 & 0.048925 & -0.188259 & 0.0875304 & 0.005066 & -0.0041743 & -0.000913 & -0.000931 & 0.0098301 & 0.0003119 & 0.0002319 & -0.002484 & -0.004193 & -0.045318 & -0.073103 & -0.045061 & 0.006121 & 0.0000509 & -0.001127 & 0.0029436 & 0.0003604 & 0.0056507 & -0.000075 & -0.000032 & 0.0000153 & 0.0000244 & -0.104771 & 0.001095 & 0.006221 & 0.0055685 & -0.000608 & -0.001711 & -0.000693 & 0.0104419 & -0.0003369 & 0.0026648 & 0.0021817 & 0.0011052 & -0.0004627 & 0.0043633 & 0.0244235 & 0.0378197 & 0.0037201 & -0.202902 & -0.000035 & -0.111647 & 0.0216544 & -0.002423 & 0.0067717 & 0.0235447 & 0.0028931 & -0.001516 & 0.0007945 & -0.005673 & -0.14797 & 0.2532717 & -0.000324 & -0.000539 & -0.000446 & 0.0133851 & 0.0116554 & -0.0007114 & -0.002348 & 0.0019199 & -0.001286 & -0.0023411 & -0.0002269 & 0.0027586 & -0.186236 & 0.0053163 & 0.0089239 & 0.0036212 & -0.027213 & -0.000308 & -0.001185 & 0.0011792 & 0.0006804 & 0.0049417 & 0.0058836 & -0.0001307 & 0.0022889 & 0.0051205 & -0.10408 & 0.002184 & 0.0429556 & 0.0033128 & 0.00010179 & -0.0005079 & 0.0001275 & -0.0002807 & -0.001317 & -0.000996 & -0.001093 & -0.0000531 & -0.0000881 & -0.000089 \end{bmatrix}$$

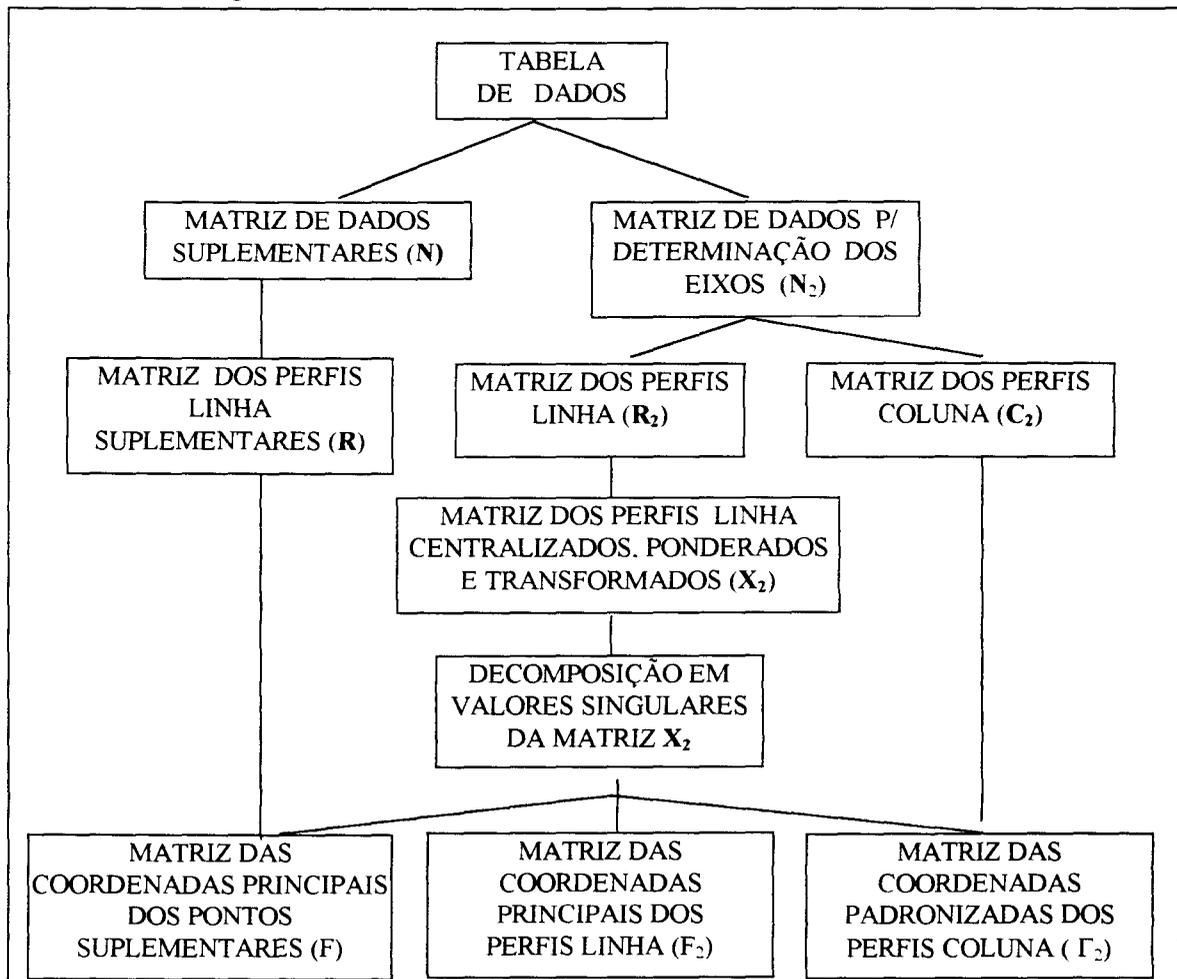
Através da expressão (2.4.23), obtemos as coordenadas padronizadas das categorias das variáveis descritoras, dadas logo abaixo em  $\Gamma_2$ . Os perfis linha estão em coordenadas principais e os perfis coluna em coordenadas padronizadas, pois nos interessa ver os perfis linha como centróides dos perfis coluna, para que possamos ver a influência (ou pesos) das variáveis descritoras nas localizações domésticas.

$\Gamma_2 =$

|           |           |           |            |           |           |            |           |           |            |            |           |           |           |
|-----------|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|------------|-----------|-----------|-----------|
| 1.436491  | 1.913417  | -0.100903 | 3.086327   | 0.2523763 | -0.561305 | 0.1293189  | 1.1010679 | 0.1324041 | -0.350543  | -0.116066  | 0.1256453 | 0.1296305 | -0.093347 |
| 1.4805296 | -0.714035 | -0.304224 | -0.810321  | -0.341336 | 0.1817655 | 0.4940397  | -0.617689 | -0.637182 | 0.3516423  | 0.1775526  | -0.043155 | -0.181423 | 0.1985852 |
| -1.46644  | -0.114717 | 0.3396618 | 0.4911513  | 0.3061337 | -0.353951 | -0.570173  | -0.649151 | -0.141801 | -0.46405   | -0.400902  | 0.1400907 | 0.186174  | -0.567574 |
| -1.121035 | -0.061211 | 0.3353602 | -0.140577  | 0.1957463 | -0.104806 | -1.007495  | 1.6188549 | 0.1951467 | 1.3563951  | 1.6210965  | -0.195503 | -0.512609 | 1.1511164 |
| 0.5453631 | 1.5317471 | 1.5346641 | -3.063402  | 0.8708495 | 0.3713564 | 0.4386863  | 0.8489111 | 0.7571445 | -1.294707  | -1.757914  | -1.031349 | 0.1321564 | -0.76209  |
| -1.105941 | -0.46741  | -7.731299 | -3.898941  | -0.812919 | 7.732376  | 1.3141851  | 1.4602294 | 17.310267 | 6.0161566  | -0.490016  | 15.764663 | -1.376016 | -7.106311 |
| 0.1601611 | 0.8121521 | 6.1978509 | -0.1119307 | 0.3111769 | 11.477561 | 3.1297186  | -7.106154 | -1.137815 | 4.614385   | 3.1008283  | 1.3347359 | -4.104967 | 0.143374  |
| -0.191366 | -1.44149  | -0.546737 | -0.391123  | 1.4021702 | -0.174499 | 0.8298485  | 0.0179946 | -1.110791 | 2.1137286  | -1.09979   | 1.4945622 | -1.774508 | -1.064962 |
| -0.022957 | -0.770708 | -0.566467 | -0.594691  | 1.4748949 | -1.815379 | 0.8026831  | -0.757468 | 1.5461454 | -0.0427    | 1.1143394  | -1.403679 | 0.0173964 | -1.931152 |
| -0.330074 | -1.061426 | -0.356056 | -1.141507  | 1.0555126 | 0.9311963 | 0.6911911  | 0.6189347 | -0.174926 | 0.5312952  | 0.1578165  | -0.931401 | 0.8390928 | 0.1536273 |
| -0.372365 | -0.663126 | -0.578208 | 0.1135003  | 0.6110344 | -0.160857 | 0.6140334  | -0.194722 | 1.3012012 | -0.043021  | -1.113613  | -0.176684 | -1.961971 | 3.495902  |
| -0.124776 | -1.205929 | -0.129746 | 0.1250065  | 0.8069033 | -0.117851 | 0.1590995  | -0.050689 | 0.1140451 | 1.4910073  | 0.7011979  | 1.0159662 | 3.0201766 | 0.8495036 |
| -0.173361 | -0.174804 | -0.531462 | -0.9217641 | -0.304435 | 0.7367829 | 0.4534711  | 0.1415863 | -1.110246 | -1.616243  | 1.7126374  | 1.8006127 | 0.9594708 | -0.492319 |
| -0.101403 | 0.0958347 | -0.565248 | 0.4094746  | -1.06338  | 1.7260611 | -1.034817  | -0.656709 | 0.0261367 | -0.802959  | 0.5043686  | -0.991947 | -1.751355 | -1.143415 |
| 0.1023491 | 0.2036131 | -0.315644 | 0.8063029  | -1.768447 | 1.5305641 | -1.063056  | 0.1216465 | 0.3538975 | 1.143145   | -1.800160  | -0.571950 | 1.6304743 | 0.1813453 |
| 0.1853351 | 1.1077664 | 0.1834359 | 0.0441479  | -1.751683 | 0.6362196 | -0.406841  | 1.1110977 | -0.193810 | 0.1516416  | 1.5483533  | -0.069647 | 0.176     | 0.2971414 |
| 0.4360202 | 0.7314637 | 0.7157961 | -1.36614   | -0.573313 | -0.451143 | -1.861661  | -0.368953 | -0.036839 | -0.3965242 | -0.448956  | 1.5016693 | -1.112375 | -0.057322 |
| -0.111179 | -1.874085 | 5.9245817 | 4.309075   | -1.857532 | -1.872872 | 0.3066552  | 0.5413309 | 1.7311535 | 0.1390687  | 0.0712437  | 1.0550897 | -0.810331 | -0.809372 |
| -0.156076 | 0.1034896 | -1.043037 | 0.112961   | -0.116381 | -0.109376 | 0.44414    | -0.057797 | -0.031533 | 0.3695605  | 0.0031109  | -0.01612  | -0.063545 | -0.004877 |
| 0.3657925 | -1.512541 | 0.3341241 | -0.861631  | 1.8005071 | 0.7379072 | -3.1033918 | 0.4024011 | 0.126234  | -0.667362  | -0.1110397 | 0.0377683 | 0.6258491 | 0.3537763 |

Através da fórmula de transição (2.5.16), obtemos as coordenadas principais das etapas de vida, como pontos suplementares. Para melhor compreender o exposto acima, é dado, no Quadro 1, o diagrama de fluxo dos dados.

QUADRO 1: Diagrama de Fluxo dos Dados



O Quadro 1 nos mostra, de modo bastante simplificado, o algoritmo da AC. Contudo, no apêndice B apresentamos uma descrição completa deste algoritmo de tal forma que se possa fazer

um programa, em qualquer linguagem computacional, para obter as informações necessárias à análise de correspondência de uma matriz de dados. Nesse apêndice ainda apresentamos um programa feito no IML/SAS, seguindo a nomenclatura dessa descrição e os resultados deste programa tendo como entrada a matriz  $N_2$ .

Depois de obtida as coordenadas dos perfis linha, dos perfis coluna e dos pontos suplementares no espaço ajustado, vamos verificar as inércias principais dos eixos para determinação do subespaço ótimo. Estas inércias são dadas pelos quadrados dos valores singulares obtidos na decomposição de  $X_2$ . Mas o PROC CORRESP do SAS nos dá uma saída apropriada para isso, que apresentamos no Quadro 2 abaixo.

QUADRO 2: Inércias Principais dos Eixos

| Valores Singulares | Inércias Principais | Percent. | 11    | 14 | 38 | 40 | 40 |
|--------------------|---------------------|----------|-------|----|----|----|----|
| 0.20077            | 0.04317             | 58.65%   | ***** |    |    |    |    |
| 0.13193            | 0.01741             | 23.65%   | ***** |    |    |    |    |
| 0.07313            | 0.00535             | 7.27%    | ***   |    |    |    |    |
| 0.05731            | 0.00329             | 4.46%    | **    |    |    |    |    |
| 0.04789            | 0.00229             | 3.12%    | *     |    |    |    |    |
| 0.03251            | 0.00106             | 1.44%    | *     |    |    |    |    |
| 0.01487            | 0.00062             | 0.84%    |       |    |    |    |    |
| 0.01366            | 0.00019             | 0.25%    |       |    |    |    |    |
| 0.01071            | 0.00012             | 0.16%    |       |    |    |    |    |
| 0.00919            | 0.00007             | 0.09%    |       |    |    |    |    |
| 0.00530            | 0.00003             | 0.04%    |       |    |    |    |    |
| 0.00335            | 0.00001             | 0.02%    |       |    |    |    |    |
| 0.00309            | 0.00001             | 0.01%    |       |    |    |    |    |
| 0.00247            | 0.00001             | 0.01%    |       |    |    |    |    |
|                    | 0.07360             |          |       |    |    |    |    |

Nesse quadro podemos observar que o primeiro eixo tem inércia de 0,04317, o que corresponde a 58,65% da inércia total e o segundo eixo tem inércia de 0,01741 correspondendo a 23,65% da inércia total, sendo a inércia total de 0,07350. O terceiro eixo representa apenas 7,27% da inércia total e, portanto, podemos considerar apenas os 2 (dois) primeiros eixos que contribuem com 82,30% da inércia total, e podem ser considerados suficientes para representar os pontos.

As coordenadas dos pontos no subespaço ótimo são dadas pelas duas primeiras colunas das matrizes  $F_2$  e  $\Gamma_2$ , apresentadas anteriormente. Representando estas coordenadas em um mesmo sistema de eixos, temos o gráfico da análise de correspondência, às vezes chamado de mapa da análise de correspondência.

Na seção seguinte, apresentamos as interpretações destes gráficos.

## **4.4 DESCRIÇÃO**

Para descrever os gráficos da Análise de Correspondência sobre a aplicação que estamos apresentando, vamos seguir o roteiro dado na seção 2.6. Deste modo, inicialmente, descrevemos os dois primeiros eixos. Para fazer esta descrição, podemos observar como se decompõe a inércia das variáveis descritoras e das localizações domésticas; essa decomposição pode ser vista no apêndice B, seção B.3. Uma outra possibilidade para a realização da descrição é observarmos na seção B.4, na saída do proc corresp, os pontos que mais contribuíram para a determinação de cada eixo e sua representatividade nestes eixos. Assim, estudamos os eixos considerando primeiramente as variáveis descritoras, já que o nosso interesse é examinar a influência destas variáveis nas localizações domésticas. Ao conseguirmos dar significado aos eixos e neles definir oposições, descrevemos as localizações domésticas sobre esta estrutura. Feito isto, estudamos a disposição dos pontos no plano formado pelos eixos considerados. Este estudo no plano tem dois momentos, primeiro, descrevemos os gráficos dos centróides dos grupos em termos das variáveis descritoras e, depois, os gráficos dos pontos suplementares são descritos em termos destas mesmas variáveis.

### **4.4.1 DESCRIÇÃO DO PRIMEIRO EIXO**

O primeiro elemento da análise deste eixo é, obviamente, a sua inércia principal, que é igual a 0,04317, o que corresponde a 58,65% da inércia total do espaço ajustado.

Para facilitar a descrição desse eixo, ou para que possamos verificar a possibilidade de dar-lhe significado, construímos o Quadro 3. Este quadro nos mostra as categorias das variáveis descritoras que mais contribuíram para a determinação do primeiro eixo, ordenadas em termos das contribuições absolutas. Ele mostra também a contribuição relativa destes pontos. Constam ainda, neste quadro, os pontos, dentre os analisados, que estão em oposição no eixo.

Assim, observando o Quadro 3, percebemos que as categorias das variáveis descritoras que mais contribuem para a determinação do eixo 1 são I2, I1, I0, Zr, I3, Rn e Zu, sendo que I0 tem contribuição absoluta (*c.a.*) ainda maior no eixo 2. Por sua vez, o Gráfico 1 nos permite

QUADRO 3: Auxílio à descrição do eixo 1.

|        | DIMENSÃO 1  |              | oposições |
|--------|-------------|--------------|-----------|
|        | <i>c.a.</i> | <i>c.r.</i>  |           |
| I2     | 0.257330    | 0.978140 (1) | I0, I1    |
| I1     | 0.253957    | 0.971121 (3) | x         |
| I0 (2) | 1.153681    | 0.643076 (6) | I2, I3    |
| Zr     | 0.142189    | 0.730367 (5) | -----     |
| I3     | 0.084666    | 0.969692 (2) | Zr        |
| Rn     | 0.059140    | 0.377742 (8) | x         |
| Zu     | 0.019239    | 0.739426 (4) | Zu        |

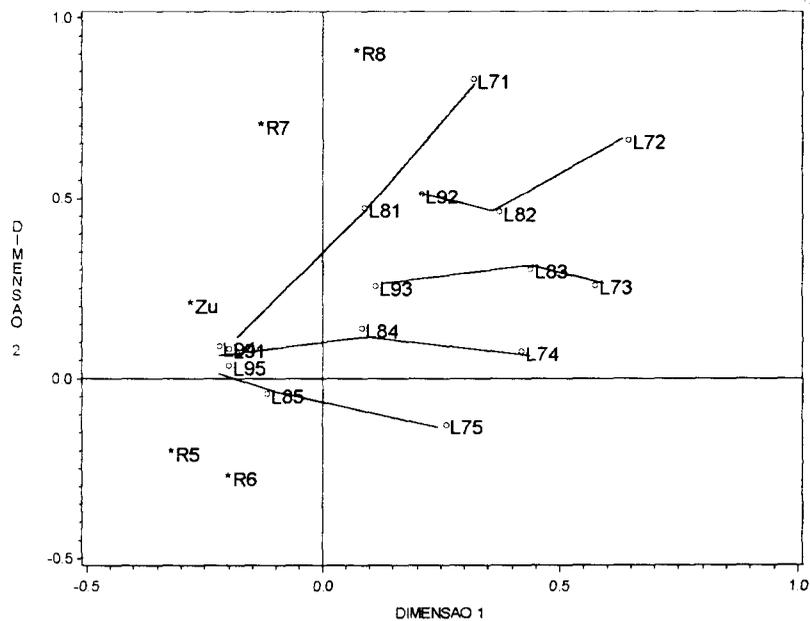
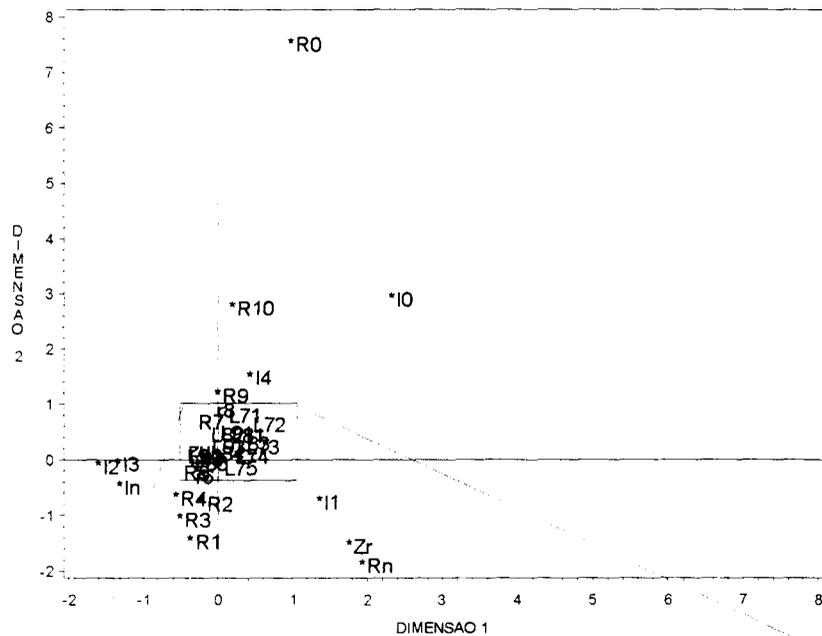
constatar que I0 e I1 estão projetadas no lado oposto a I2 e I3; e que Zr também tem sua projeção no lado oposto ao de Zu. Ao estudar a qualidade de representação destes pontos, observamos que todos eles estão bem representados, isto é, que todos têm contribuição relativa alta.

O número que aparece à direita na coluna das contribuições relativas (*c.r.*), no Quadro 3 acima, indica a posição da categoria no rol destas contribuições. Logo, podemos dizer que o eixo 1, ou dimensão 1, caracteriza a instrução dos indivíduos da população e a localização de seu domicílio. Desta forma, os indivíduos das categorias projetadas no lado direito (+) do eixo tendem a localizar-se, em maior proporção, na zona rural e a ter instrução com menos anos de estudo. Por sua vez, aquelas categorias projetadas no lado esquerdo (-) tendem a ter mais anos de estudo - não mais que onze anos- e a ter uma maior proporção na zona urbana.

Tendo dado significado ao primeiro eixo, podemos estudar, sob esta estrutura, as localizações domésticas. Agora, segundo Greenacre (1984) -“apelando para nossa habilidade visual de observar imagens e esperando ter uma visão global da informação, para que possamos estimular possíveis explicações”- passamos a examinar as localizações domésticas ao longo dos três momentos do tempo, isto é, nos censos demográficos de 1970, 1980 e 1991.

Focalizando o “zoom” do Gráfico 1, podemos distinguir que as categorias da variável localização doméstica se projetam da direita para a esquerda, ao longo dos censos. Isto é, que todas as localizações domésticas se projetam neste sentido, primeiro as do censo de 1970, depois as de 1980 e 1991. Os pontos nos três censos das localizações L1, L4 e L5 estão projetados mais à esquerda (-) que os de L2 e L3. Assim, podemos concluir que as pessoas que vivem sós e as que são chefes de família monoparental formam grupos com menor instrução na população.

GRÁFICO 1: Mapa da AC dos Centróides dos Grupos e as variáveis descritoras



LEGENDA

LOCALIZAÇÃO DOMÉSTICA

- Lc1=vive em grupos de pessoas sem laços de parentesco
- Lc2=vive só
- Lc3=vive como chefe de família monoparental
- Lc4=vive em grupos aparentados
- Lc5=vive em família chefiada por outros

CENSO DEMOGRÁFICO

c=7,8 e 9(1970,1980 e 1991)

INSTRUÇÃO

- In=instrução não declarada
- I0=sem instrução
- I1=de 1 a 4 anos de estudo
- I2=de 5 a 8 anos de estudo
- I3=de 9 a 11 anos de estudo
- I4=de 12 ou + anos de estudo

ZONA

- Zu=zona urbana
- Zr=zona rural

RENDA

- Rn=renda não declarada
- R0=sem renda
- R1=renda do 1 decil
- R2=renda do 2 decil
- R3=renda do 3 decil
- R4=renda do 4 decil
- R5=renda do 5 decil
- R6=renda do 6 decil
- R7=renda do 7 decil
- R8=renda do 8 decil
- R9=renda do 9 decil
- R10=renda do 10 decil

Oliveira, M. C. F. A, nos revela que estudos populacionais têm mostrado uma ampla urbanização da população a partir de 1960. Porém, o gráfico 1 mostra que o grupo de pessoas que vivem sós e o dos chefes de família monoparental são menos urbanizados. Isto mostra a existência de uma forte associação entre estes tipos de localizações domésticas e a residência rural, o que não significa que a maior proporção dos que vivem sós e dos que são chefes de família monoparental esta na zona rural. Também, podemos dizer que existe uma forte associação entre viver só, ser chefe de família monoparental e ter nível de instrução baixa.

#### 4.4.2 DESCRIÇÃO DO SEGUNDO EIXO

Este eixo contribui com 23,65% da inércia total e as variáveis descritoras que mais contribuem para a sua formação são I0, R10, R0, Zr, I1, R1, R9, I4, R3, R8, R2, R7 e R4, como pode ser observado no Quadro 4. Embora Zr e I1 tenham altas contribuições absolutas na determinação da direção deste eixo, elas são maiores no eixo 1. Observando agora o Gráfico 1, percebemos que as categorias de renda estão dispostas em ordem crescente de baixo para cima. Desta forma, temos a oposição entre as renda menores (R1, R2, R3 e R4) em baixo e maiores (R7, R8, R9 e R10) em cima, já que estas são as categorias que estão melhor representadas neste eixo, como apresentamos no quadro abaixo. Assim, a dimensão 2 -eixo 2- é caracterizada pela variável renda.

QUADRO 4: Auxílio à descrição do eixo 2.

| DIMENSÃO 2 |             |               |           |
|------------|-------------|---------------|-----------|
|            | <i>c.a.</i> | <i>c.r.</i>   | oposições |
| I0         | 0.220982    | 0.314728 (11) | R1,R2,    |
| R10        | 1.184601    | 0.724844 (2)  | R3,R4     |
| R0         | 1.114946    | 0.547030 (5)  | x         |
| Zr (1)     | 1.092686    | 0.191966 (13) | R7,R8,    |
| I1 (1)     | 1.063939    | 0.088469 (17) | R9,R10    |
| R1         | 1.063541    | 0.774555 (1)  |           |
| R9         | 1.040467    | 0.699654 (3)  | -----     |
| I4         | 0.040218    | 0.315362 (10) |           |
| R3         | 1.039813    | 0.698823 (4)  |           |
| R8         | 1.026070    | 0.519611 (6)  |           |
| R2         | 0.018696    | 0.481835 (7)  |           |
| R7         | 1.016340    | 0.464121 (8)  |           |
| R4         | 1.016097    | 0.458355 (9)  |           |

Considerando as variáveis descritoras nesse eixo e observando as localizações domésticas, percebemos que as localizações L5, L4 e L3, consideradas cada uma ao longo dos censos, estão dispostas em padrão de paralelismo e em deslocamento na direção de menores para maiores rendas. Ou seja, cada uma destas localizações, ao longo dos censos, estão associadas a patamares de renda semelhantes. A localização L2 apresenta um padrão diferente ao longo dos censos: o movimento entre o censo de 1970 e o de 1980 se dá no sentido contrário e com amplitude maior que a das localizações domésticas anteriormente comentadas. A L1 se dispõe mais transversal no gráfico, isto é, cai o patamar de renda desta categoria, já que os indivíduos que vivem sem laços de parentesco, em 1970, tendem a ter renda mais alta do que os indivíduos que compõem esta categoria em 1991. Assim, o padrão de quem vive em grupo de pessoas sem laços de parentesco em 1970 não é o mesmo que em 1991.

Observamos ainda que a renda discrimina as localizações em todos os momentos, agrupando-as de maneiras diferentes, mas não modificando o tipo de influência que exerce ao longo dos tempos. Isto é, a renda distingue mais o tipo de localização doméstica, do que os censos dentro de uma mesma localização.

Descritos os dois eixos, vamos agora descrever a posição dos pontos no plano.

#### 4.4.3 A DESCRIÇÃO DO GRÁFICO DOS CENTRÓIDES

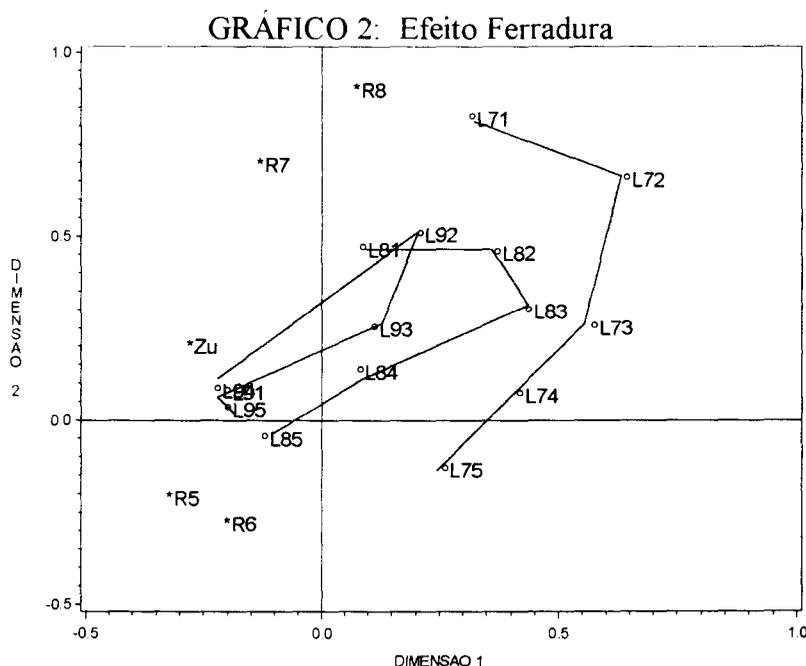
Estudando o Gráfico 1, percebemos que as variáveis descritoras estão mais distantes da origem do sistema. Isso ocorre porque elas estão representadas em coordenadas padronizadas, enquanto as localizações domésticas estão em coordenadas principais. Olhemos, com mais detalhe, a nuvem das variáveis descritoras.

Percebemos, que os pontos Zr e Rn estão próximos entre si, observando a matriz de perfis coluna, percebemos que seus perfis são semelhante. Assim, somado ao fato da contribuição relativa destes pontos ser mais ou menos próxima, podemos dizer que Zr e Rn estão associados de forma semelhante a todas as localizações domésticas. Isso sugere que a maior proporção de rendas não declaradas refere-se à zona rural. Uma outra observação que podemos fazer, neste gráfico, é que o ponto R0 está muito distante da origem. Isto se deve à sua pouca massa, que corresponde a 0,2% da massa total da nuvem. O ponto de maior massa é o Zu com 29,33% da massa total da

nuvem, seguida do I2 com 11,64% e I1 com 11,55% da massa total.

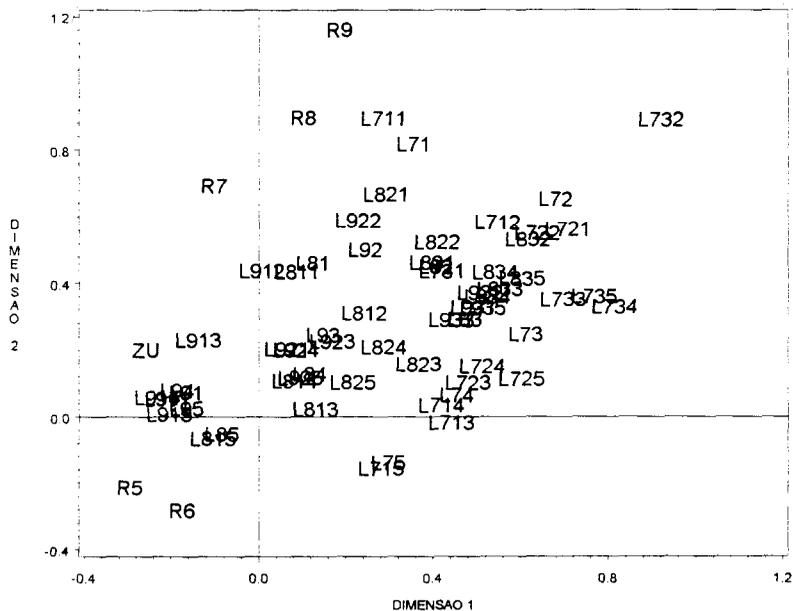
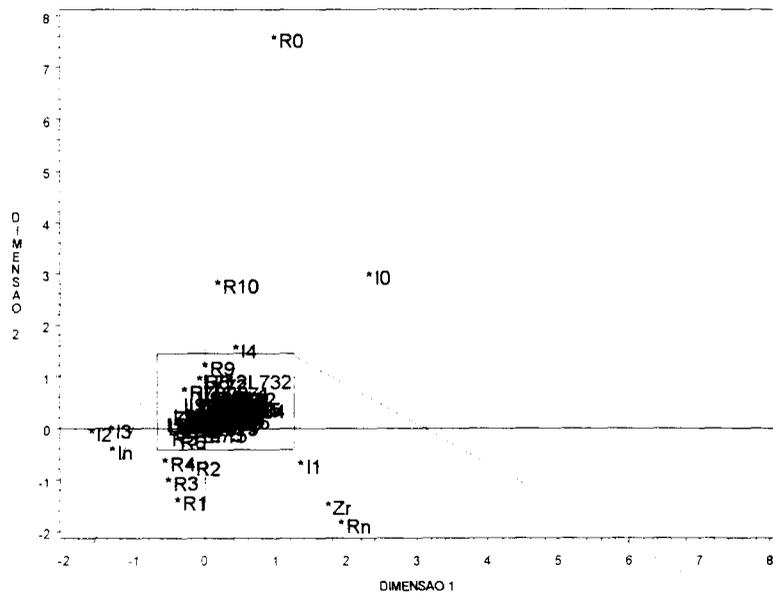
O fato de uma nuvem estar em coordenadas principais e a outra em coordenadas padronizadas, nos possibilita ver as localizações domésticas como centróides (baricentros) das variáveis descritoras, como dadas pela fórmula de transição descrita em (2.6.8). Assim podemos perceber a influência destas variáveis sobre as localizações domésticas.

No nosso caso, é interessante notar que a renda e a instrução se relacionam de forma quadrática na sua influência sobre as localizações domésticas. Esta relação é mais patente no censo de 70, como pode ser observada no gráfico 2. Ela também ocorre no censo de 80, porém com uma curva mais fechada. No censo de 91, a relação quadrática da renda e instrução na influência sobre as localizações domésticas se fecha. Esta relação, discutida em Van Rijckevorsel (1986), é denominada **efeito ferradura**.



No Gráfico 2 é possível perceber que, inicialmente, renda e instrução tinham um tipo de interação. Esta interação foi alterada ao mudar de censo, com a modificação no nível de instrução e renda da população. Na verdade, a curva vai se fechando para mostrar a diminuição da diferença das localizações domésticas ao longo do tempo. Tomando os extremos da ferradura e observando ao longo dos censos, vemos que L5 move-se de menor para maior instrução e renda; já L1 move-se de menor para maior instrução e de maior para menor renda.

GRÁFICO 3: Mapa da AC com os Pontos Suplementares incluídos



## LEGENDA:

## LOCALIZAÇÃO DOMÉSTICA

Lce1=vive em grupos de pessoas  
sem laços de parentesco  
Lce2=vive só  
Lce3=vive como chefe de família  
monoparental  
Lce4=vive em grupos aparentados  
Lce5=vive em família unificada  
por outros

## CENSO DEMOGRÁFICO

c=7, 8 e 9 (1970, 1980 e 1991)

## ETAPAS DO CICLO DE VIDA

e=1, 2 e 3 (1ª, 2ª e 3ª etapa de vida)

## INSTRUÇÃO

Ir=instrução não declarada  
I0=sem instrução  
I1=instrução de 1 a 4 anos  
I2=instrução de 5 a 8 anos  
I3=instrução de 9 a 11 anos  
I4=instrução de 12 ou + anos

## ZONA

Zu=zona urbana  
Zr=zona rural

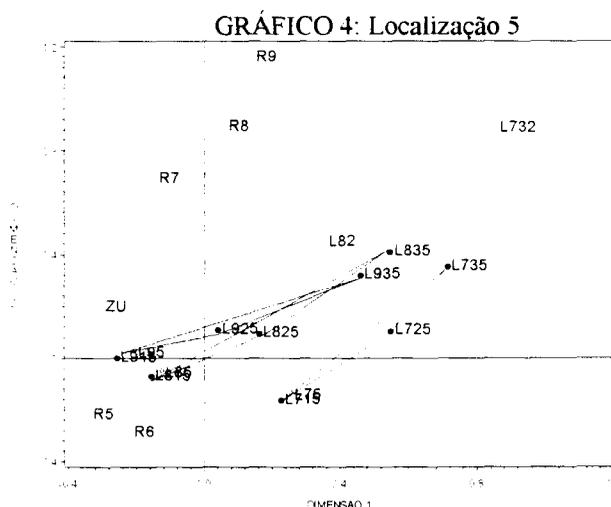
## RENDA

Rn=renda não declarada  
R0=sem renda  
R1=renda do 1 decil  
R2=renda do 2 decil  
R3=renda do 3 decil  
R4=renda do 4 decil  
R5=renda do 5 decil  
R6=renda do 6 decil  
R7=renda do 7 decil  
R8=renda do 8 decil  
R9=renda do 9 decil  
R10=renda do 10 decil

#### 4.4.4 DESCRIÇÃO DOS PONTOS SUPLEMENTARES

Observado o comportamento da localização doméstica em termos das variáveis descritoras entre elas e ao longo do tempo, vamos agora considerar uma partição da população dentro de cada localização. Ou, de outro modo, vamos considerar as etapas de vida em cada localização, o que nos permite observar o comportamento dos indivíduos com relação às etapas de vida. Lembramos que a **1ª etapa de vida (E1)** refere-se aos indivíduos *jovens solteiros*, a **2ª etapa de vida (E2)** aos *maduros separados ou divorciados* e a **3ª etapa de vida (E3)** aos *idosos viúvos*.

Ao colocar nos gráficos as etapas de vida, tomamos como referência os seus centróides, no caso, as localizações domésticas, como podemos observar no Gráfico 3. Mas observar essa imagem visual neste gráfico, mesmo em “zoom”, é uma tarefa difícil. Para melhorar as condições de observação, apresentamos a parte em “zoom” do Gráfico 3 em cinco gráficos, cada um considerando apenas uma localização e seus pontos suplementares nos três censos.



No Gráfico 4, apresentamos as etapas do ciclo de vida que se referem aos indivíduos que vivem em famílias chefiadas por outros. Neles, podemos observar que a norma da localização 5 são os indivíduos da primeira etapa de vida. Estes indivíduos representam os jovens solteiros que correspondem a 95% daqueles que vivem em família chefiada por outros, e 87% de toda a população em estudo.

O afastamento com relação à norma se dá ao longo dos três momentos do ciclo de vida. Isto é verdade em cada um dos censos, ou seja, em 1970, 1980 e 1991. Este afastamento parece resultar da influência do ponto I0 (Gráfico 3). Ou seja, a ausência de instrução associada à vida

daqueles que vivem em família chefiada por outros, afeta mais especialmente o grupo de idosos, e menos o grupo de maduros. Os jovens, isto é, a norma, é pouco afetada.

Outra observação, é que as três etapas de vida encontram-se quase em linha reta transversal no gráfico. Isto significa que além do nível de instrução estar associado às etapas de vida, a renda também tem uma relação com estas etapas. Mas esta relação tem um sentido oposto, visto que os indivíduos mais velhos tendem a ter mais renda e menos instrução e os indivíduos mais jovens tendem a ter menos renda e mais instrução.

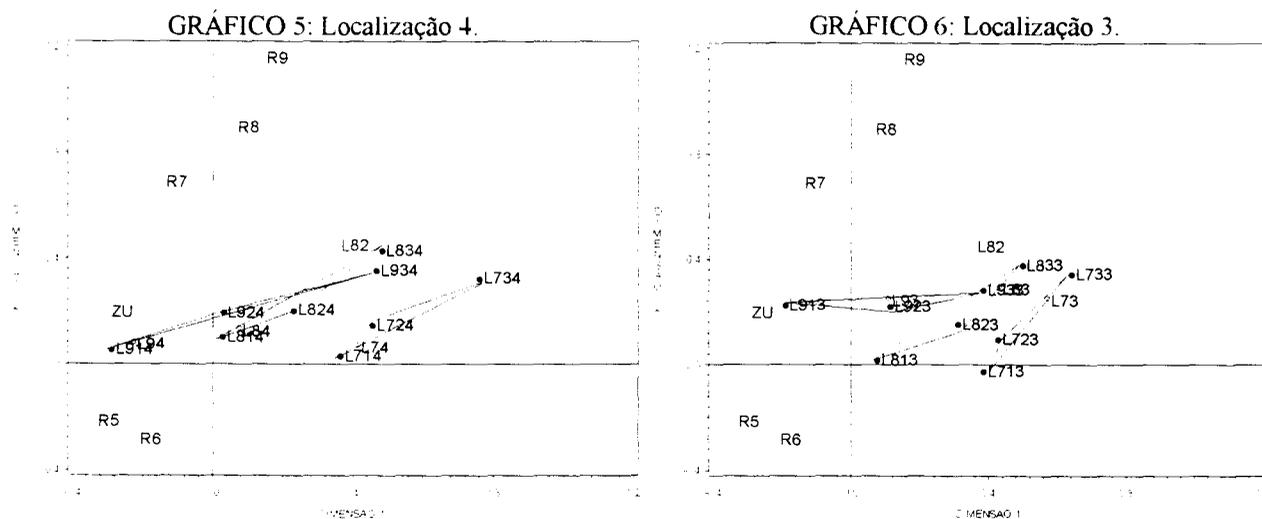
O que se observa no Gráfico 4 é que ao longo dos censos o diferencial de renda nas etapas de vida parece diminuir de 1970 para 1991. Em contrapartida, a instrução move-se em direção aos maiores níveis, dos censos de 1970 para 1991. Como já havíamos dito isto é, provavelmente, decorrente da ampliação do patamar de instrução da população brasileira. Portanto, considerando a ordem entre as etapas, os mais jovens tendem a ter mais instrução e menos renda, comparativamente aos indivíduos das outras etapas de vida.

Como podemos observar, a nuvem dos pontos suplementares referentes ao censo de 1970 é quase que paralela à dos pontos referentes ao censo de 1980. Por sua vez, esta nuvem é cortada por aquela formada pelos pontos referentes ao censo de 1991. Isto significa que houve uma modificação no nível de renda dos indivíduos idosos e dos jovens - a renda dos idosos, de 1980 para 1991, tende a diminuir e a dos jovens a aumentar. Os indivíduos da segunda etapa de vida quase não têm variação na renda.

Ao longo dos censos, observamos que os pontos relativos aos jovens vão se aproximando do ponto referente à zona urbana e dos pontos das rendas acima do sétimo percentil. Isto pode estar refletindo uma tendência no País, pois os jovens solteiros têm uma tendência a serem mais instruídos e a serem atraídos pela zona urbana e pelas rendas mais elevadas. Por sua vez, os idosos viúvos de 1991 são mais instruídos que os dos censos anteriores, contudo, apesar disto, houve uma queda na renda.

No Gráfico 5, colocamos em “zoom” a localização L4, dos indivíduos que vivem em grupos aparentados. Este gráfico tem o mesmo aspecto do gráfico anterior, e sua norma também é a categoria dos jovens solteiros. Nele podemos observar, mais uma vez que o movimento, ao longo das etapas de vidas e ao longo dos censos, é mais visível em direção às instruções mais altas. Observa-se também uma semelhança com os indivíduos que vivem em famílias chefiadas por

outros. Contudo, os jovens que vivem em grupos aparentados apresentam um aumento no poder aquisitivo nos três censos.

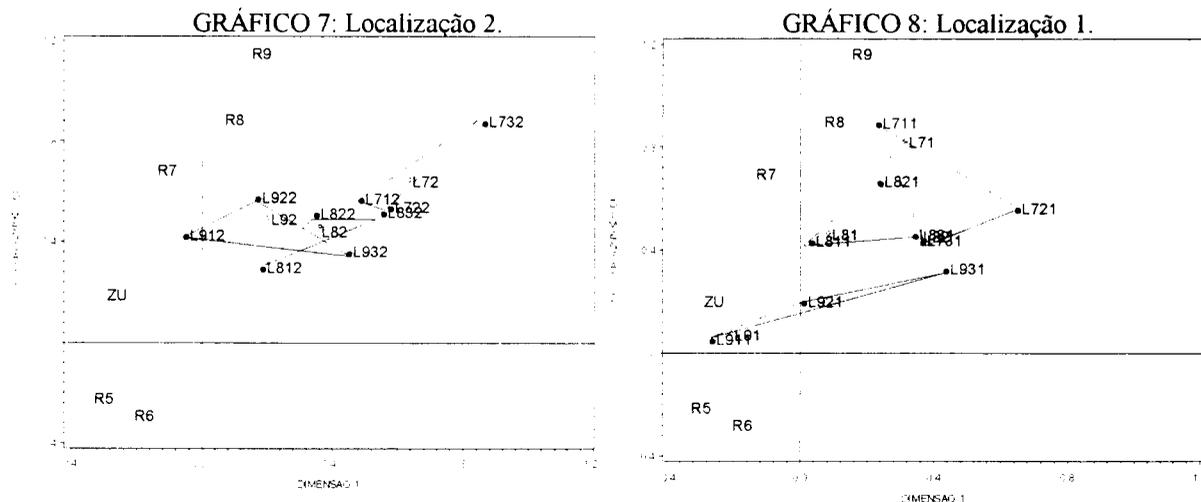


O Gráfico 6 apresenta o “zoom” da localização 3, ou seja, dos indivíduos que chefiam famílias monoparentais. Não se observa uma predominância deste tipo de localização doméstica em nenhuma das etapas do ciclo de vida, embora haja uma concentração relativamente maior e equilibrada na segunda e terceira etapas, no que se refere aos indivíduos maduros e idosos, especialmente no censo de 1970. No censo de 1991, a predominância relativa está na segunda etapa do ciclo de vida, ou seja, parece haver, ao longo dos censos, um rejuvenescimento dos indivíduos que chefiam famílias monoparentais.

Ao que parece, neste censo, os homens que chefiam famílias monoparentais têm alguma semelhança com aqueles que vivem em famílias chefiadas por outros ou em grupos aparentados; embora eles sejam mais urbanos que os demais e tendam a ter rendas superiores -pelo menos os mais jovens e os maduros. Com relação aos idosos, em todos os censos, o comportamento é exatamente igual ao das localizações anteriores. Isto quer dizer que o homem chefe de família monoparental -além de não optar, escolher, ou ser levado a morar em grupos aparentados ou morar em família chefiada por outros- têm filhos, são mais urbanos e têm condições financeiras melhores. Nos censos de 1970 e 1980, observa-se o mesmo comportamento das duas localizações anteriores.

O Gráfico 7 coloca em “zoom” a categoria dos indivíduos que vivem sós, ou seja, a localização 2. Nele percebemos a ausência da norma, já que as etapas com maiores proporções estão entre os dois ciclos de vida mais altos. Nesta categoria, em cada censo, há um arranjo diferente quanto à influência das variáveis descritoras. Assim, no censo de 1970, os indivíduos do

primeiro e segundo ciclo estão próximos, se opondo aos do terceiro. Já no censo de 80, a norma é o segundo ciclo, sendo que os indivíduos do primeiro e terceiro ciclo estão em posições opostas. No censo de 91, há uma mudança na posição em relação a renda e instrução: os indivíduos do terceiro ciclo tendem a ter menos renda e menos instrução que os do primeiro e segundo ciclo. Esta é uma situação contrária à dos censos anteriores para os indivíduos que vivem sós.



O Gráfico 8, da localização 1, do grupo de indivíduos sem laços de parentesco, apresenta outro padrão. Aqui a norma são os indivíduos do primeiro ciclo e há uma diferenciação maior dos três censos pela renda. O de 1991 tem a diferenciação, entre as etapas de vida, realçada principalmente pela instrução, a renda tem uma participação menor. Quanto à instrução, vemos um acréscimo linear à medida que evolui o ciclo de vida. Em 1980 e 1970, há uma influência maior da variável renda na diferenciação entre as três etapas de vida. Em termos de instrução, a ordem é alterada no censo de 1970. Então, temos os indivíduos do primeiro ciclo com maior renda e maior instrução, os do segundo ciclo com renda intermediária e menor instrução e os do terceiro ciclo com menor renda e instrução maior que o do segundo. As pessoas que vivem em grupos de pessoas sem laço de parentesco, em 1970 são predominantemente jovens, nos três censos em estudo.

No Gráfico 8, no censo de 1970, é interessante observar que os jovens têm menos instrução - o que é compatível com o fato de serem jovens-, mas têm renda alta. Comparativamente, estes jovens que vivem em grupos sem laços de parentesco parecem ter condições de instrução piores do que os outros jovens; eles estão atraídos pelo ponto I0. Ao longo dos três censos, os jovens: em 1970 têm maior renda e maior instrução; em 1980, têm maior instrução e renda comparável aos idosos, mas menor que os do maduros: em 1991, têm menor

renda e maior instrução.

Se compararmos este gráfico com os demais, notamos que os pontos da localização 1 e da localização 2 estão bem acima do eixo renda; eles representam os grupos de indivíduos que tendem a estar num patamar de renda melhor do que os outros. Isto faz sentido, pois segundo Oliveira M. C. F. A. (comunicação pessoal), o fato de se ter maior independência econômica permite criar arranjos domésticos por conta própria, sem estar vinculado a laços de parentesco que podem implicar um mecanismo de redistribuição de recursos. Segundo ela, a renda parece estar associada a uma maior ou menor independência em relação ao mecanismo de proteção social que é a família.

Essa melhoria da renda das pessoas que vivem sós e em grupos não aparentados, nas diversas etapas de vida e nos diversos censos destes grupos, se altera em 1991. O censo deste ano reflete o padrão que se observa nas outras localizações domésticas. O grupo dos indivíduos viúvos idosos tem uma renda ligeiramente maior e, certamente, pior instrução. O grupo das pessoas maduras e divorciadas ou separadas tem melhor condição de instrução, mas renda relativamente menor que a dos idosos viúvos. De fato, segundo a pesquisadora acima citada, isto pode ser efeito do ciclo de vida profissional, até porque as pessoas maduras e divorciadas ou separadas formam um grupo grande e heterogêneo.

Os jovens são a norma e, em relação aos demais indivíduos do mesmo censo, têm melhores condições de instrução e menor renda. Não que sua renda seja muito ruim, pois eles são atraídos pelas rendas intermediárias (R5 e R6), mas, de qualquer forma, afastam-se de R7 e R8. Assim, eles parecem formar uma outra população, ou seja, de censo para censo os jovens mudam. Para Oliveira, provavelmente, isto está relacionado a processos em curso no País. Segundo ela, é possível que esta mudança esteja relacionada a alterações de aspecto educacional, ou a mudanças na distribuição de renda do País -porque há uma piora relativa nos patamares de renda- e é possível ainda que esteja relacionada a políticas habitacionais que permitiram às pessoas fazerem arranjos domésticos independentes. Oliveira acredita que mudanças deste tipo podem estar alterando o perfil das pessoas que escolhem, ou são levadas a morar em grupos sem laços de parentesco. Deste modo, observamos no Gráfico 8, que os jovens foram os diferenciadores do comportamento dos indivíduos ao longo dos censos.

Os gráficos 7 e 8 -que contemplam, respectivamente, a categoria dos indivíduos que vivem sós e a dos não aparentados- parecem apresentar um padrão completamente diferente dos outros

três. Em primeiro lugar, não se observa nestes gráficos, o paralelismo entre os feixes constituídos pelas três etapas de ciclo de vida, em cada um dos censos. Em segundo lugar, no senso de 1970, em ambos os gráficos, os indivíduos tendem a rendas mais altas. Analisando o grupo dos idosos ao longo dos três censos, e sempre em relação aos seus companheiros de outras etapas no mesmo censo, temos que eles têm menor instrução e maior renda. No censo de 1980, os idosos continuam a ter menor instrução, porém a renda é equivalente à dos indivíduos da segunda etapa de vida. Em 1991 eles têm menor instrução e menor renda.

Tanto a localização 2 quanto a localização 1 apresentam padrões diferentes entre si e das demais localizações; elas têm as formas de arranjo que mais se afastam do conceito de família. Aqueles que vivem em família chefiada por outros, são filhos, agregados, sogros ou pais de um dos membros do casal. Os que vivem em grupos aparentados, que não tem como núcleo central o casal, também são pessoas que mantêm relações de parentesco. Os chefes de família monoparental se aproximam mais de um modelo de família -alguns até chamam de família incompleta- mas é um tipo de família que está crescendo no Brasil. Por sua vez, os indivíduos das localizações 2 e 1 são pessoas que, ou estão vivendo isoladas de um núcleo familiar, vivendo sós, ou vivem com outras pessoas com as quais não têm relações de parentesco. Os dados mostram que, pelo menos no caso dos homens, as pessoas que não vivem em arranjos domésticos do tipo familiar, ou que envolvem relações de parentesco, são diferentes daquelas que vivem em arranjos domésticos que se aproximam de alguma forma da noção de família, por se manterem em relação de parentesco.

As localizações 5, 4 e 3 revelam uma ordem ao longo das etapas de vida: os mais jovens tendem a uma maior instrução e menor renda. Ressalta-se que no censo de 1991, na localização 3, a influência da renda não parece estar presente e a ordem das etapas apresenta-se ao longo das instruções. O diferencial, ao longo dos censos, é acentuado devido à instrução.

Finda a descrição dos gráficos, acreditamos que a técnica utilizada foi capaz de fornecer várias e interessantes informações, mostrando-se, realmente, uma poderosa ferramenta para a análise dos dados. Deste modo, acreditamos que, por um lado, os dados foram capazes de levar-nos a uma melhor compreensão da utilização prática da Análise de Correspondência em Grupo. Por outro lado, sem dúvida, as informações aqui apresentadas constituem uma parte importante do estudo sobre a localização doméstica dos homens sem companheiras no Estado de São Paulo. É claro que estas informações serão posteriormente analisadas por pesquisadores da área de estudos populacionais.

## 5. CONSIDERAÇÕES FINAIS

### 5.1 SOBRE O TRABALHO

O objetivo deste trabalho foi o de *apresentar o desenvolvimento a Análise de Correspondência para dados que tenham estrutura de grupo, e cujo interesse fosse realçar a diferenciação entre os grupos*. Para tanto, utilizamos especialmente as idéias de M. J. Greenacre. No capítulo 2, mostramos o desenvolvimento do algoritmo da Análise de Correspondência para tabelas de contingência onde não se considera a estrutura de grupo. Foi sobre este pilar que, no capítulo 3, apresentamos a Análise de Correspondência para dados com estrutura de grupo.

Contudo, não bastava desenvolver a técnica, era necessário experienciar seu poder no trato de uma tabela mais complexa. Isto foi feito no capítulo 4, por meio do estudo sobre as localizações domésticas de homens sem companheiras do Estado de São Paulo.

Obviamente, ao longo deste trabalho, não colocamos idéias originais, visto que em Greenacre (1984) já se vislumbra a origem da Análise de Correspondência em Grupo, e em Greenacre e Blasius (1994) já se discute a aplicação de tabelas concatenadas. Contudo, tanto estas como outras publicações, não desenvolvem algebricamente a técnica, a ponto de um estatístico tomar posse e intimidade com seu conteúdo. Por esta razão, tentamos colocar, ao alcance do estatístico o desenvolvimento algébrico detalhado das idéias plantadas em 1984. Estas idéias eram privilégio daqueles que estavam familiarizados com a Análise de Correspondência Simples e que passavam para a Análise de Correspondência em Grupos sem a explicitação da estrutura inerente a esta nova visão e sem detalhar algébrica e geometricamente.

Mas o principiante da Análise de Correspondência não tem acesso a esta visão. Então ele teria que, praticamente, percorrer o caminho das aplicações para adquirir esta sensibilidade. Este trabalho permite encurtar o caminho de intimidade com a técnica, do seu uso diferenciado em situações mais complexas de sobre-estrutura, tanto de linhas quanto de colunas. Neste sentido justifica-se o nível do detalhamento algébrico e a presença dos apêndices; estes últimos objetivando um possível acompanhamento detalhado da análise, bem como facilitando ao leitor interessado a reexecução destas e a execução de novas análises.

## **5.2 SOBRE A TÉCNICA**

A técnica de Análise de Correspondência, como uma técnica de dados multivariados, permite que os dados falem por si só, sem uma hipótese de modelagem ou uma distribuição de probabilidade. Neste contexto, esta técnica pode ser suficiente em si mesma, como fizemos neste trabalho, ou pode ser complementar a outros tipos de modelagem. Para entender o significado de um modelo pode ser necessário o auxílio gráfico ou, ao contrário, uma visão gráfica pode levar à necessidade de se criarem hipóteses de modelos para serem confirmadas. Estas duas maneiras de tratar os dados têm uma dinâmica de comunicação.

A assimilação da estrutura de grupo modifica a ótica do objetivo da própria análise, que passa a ser não uma associação de linhas e de colunas, mas sim o entendimento dos padrões de diferenciação ao longo dos grupos, seja por comportamento linha, seja por comportamento coluna. Daí, podemos dizer que, na realidade, a análise de correspondência em grupos é uma Análise Discriminante para dados categóricos.

Finalmente, gostaríamos ainda de lembrar que a Análise de Correspondência em Grupo é desenvolvida em três etapas: a) definir um representante para cada grupo, b) maximizar a inércia destes representantes e c) colocar todas as categorias como pontos suplementares. Em Análise Discriminante há estudos sobre a definição ou escolha de um representante ideal para os grupos (Takane 1989). Fica, portanto, como sugestão de investigações futuras, a implementação desta ideia para o caso da Análise da Correspondência em Grupos.

## APÊNDICE A

**A.1 ACS DA MATRIZ N**

Aqui apresentaremos conceitos e resultados análogos aos do estudo feitos no capítulo 2. Mas, considerando a estrutura de grupos na matriz de dados, as notações para os seus elementos serão modificadas e eles terão dois novos índices que representarão os grupos.

Portanto a matriz de dados,

$$N = \begin{bmatrix} n_{11}^{11} & \cdots & n_{1J_1}^{11} & \cdots & n_{11}^{1L} & \cdots & n_{1J_L}^{1L} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ n_{I_1 1}^{11} & \cdots & n_{I_1 J_1}^{11} & \cdots & n_{I_1 1}^{1L} & \cdots & n_{I_1 J_L}^{1L} \\ \vdots & & \vdots & & \vdots & & \vdots \\ n_{11}^{H1} & \cdots & n_{1J_1}^{H1} & \cdots & n_{11}^{HL} & \cdots & n_{1J_L}^{HL} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ n_{I_H 1}^{H1} & \cdots & n_{I_H J_1}^{H1} & \cdots & n_{I_H 1}^{HL} & \cdots & n_{I_H J_L}^{HL} \end{bmatrix}$$

Matriz de Correspondência

$$P = 1/nN$$

enquanto o vetor de massas dos perfis linhas e centroide dos perfis colunas serão denotados por

$$r = \left( \frac{n_{1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{I_1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{1\bullet}^{H\bullet}}{n}, \dots, \frac{n_{I_H\bullet}^{H\bullet}}{n} \right),$$

Por sua vez, o vetor de massas dos perfis colunas e centróide dos perfis linhas serão dados como

$$c = \left( \frac{n_{\bullet 1}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet J_1}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet 1}^{\bullet L}}{n}, \dots, \frac{n_{\bullet J_L}^{\bullet L}}{n} \right),$$

e o i-ésimo perfil linha do h-ésimo grupo por

$$R_i^h = \left( \frac{n_{i1}^{h1}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{iJ_1}^{h1}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{i1}^{hL}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{iJ_L}^{hL}}{n_{i\bullet}^{h\bullet}} \right), \quad i = 1, \dots, I_h \text{ e } h = 1, \dots, H.$$

A notação utilizada para a matriz dos perfis linhas será

$$R = [(R_1^1)' \dots (R_{I_1}^1)' \dots (R_1^H)' \dots (R_{I_H}^H)']'$$

e para o j-ésimo perfil coluna do ℓ-ésimo grupo, teremos

$$C_j^\ell = \left( \frac{n_{1j}^{1\ell}}{n_{\bullet j}^{\bullet \ell}}, \dots, \frac{n_{I_1 j}^{1\ell}}{n_{\bullet j}^{\bullet \ell}}, \dots, \frac{n_{1j}^{H1}}{n_{\bullet j}^{\bullet \ell}}, \dots, \frac{n_{I_H j}^{H\ell}}{n_{\bullet j}^{\bullet \ell}} \right), \quad j = 1, \dots, J_\ell \text{ e } \ell = 1, \dots, L.$$

Finalmente, a matriz dos perfis colunas será colocada como

$$C = [(C_1^1)' \dots (C_{J_1}^1)' \dots (C_1^L)' \dots (C_{J_L}^L)']'$$

A distância definida para a nuvem dos pontos-linhas no espaço  $\mathbf{R}^J$  é dada pela distancia euclidiana generalizada, definida pela matriz diagonal  $D_c$ , onde na diagonal temos os elementos do centróide, é dada por

$$\begin{aligned} d((R_i^h), (R_{i'}^{h'})) &= \sum_{\ell=1}^L \sum_{j=1}^{J_\ell} \left( \frac{n_{ij}^{h\ell}}{n_{i\bullet}^{h\bullet}} - \frac{n_{i'j}^{h'\ell}}{n_{i'\bullet}^{h'\bullet}} \right)^2 \frac{n}{n_{i\bullet}^{h\bullet}} \\ &= ((R_i^h) - (R_{i'}^{h'}))' D_c^{-1} ((R_i^h) - (R_{i'}^{h'})) \end{aligned}$$

e distância definida para a nuvem dos pontos-colunas no espaço  $\mathbf{R}^I$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_r$ , onde na diagonal temos os elementos do centróide, é dada por

$$\begin{aligned} d((C_j^i), (C_{j'}^{i'})) &= \sum_{h=1}^H \sum_{\ell=1}^{I_h} \left( \frac{n_{ij}^{h\ell}}{n_{\bullet j}^{h\bullet}} - \frac{n_{i'j'}^{h'\ell}}{n_{\bullet j'}^{h'\bullet}} \right)^2 \frac{n}{n_{i\bullet}^{h\bullet}} \\ &= ((C_j^i) - (C_{j'}^{i'}))' D_r^{-1} ((C_j^i) - (C_{j'}^{i'})) \end{aligned}$$

A Matriz que define o ajuste da nuvem do espaço  $\mathbf{R}^J$  é dada por

$$\mathbf{S} = (\mathbf{P} - \mathbf{rc}')' D_r^{-1} (\mathbf{P} - \mathbf{rc}') D_c^{-1}$$

$$\mathbf{S} = \mathbf{X}' \mathbf{X} \text{ onde}$$

$$\mathbf{X} = D_r^{-1/2} (\mathbf{P} - \mathbf{rc}') D_c^{-1/2}$$

e a Matriz que define o ajuste da nuvem do espaço  $\mathbf{R}^I$  é dada por

$$\mathbf{Z} = (\mathbf{P}' - \mathbf{cr}')' D_c^{-1} (\mathbf{P}' - \mathbf{cr}') D_r^{-1}$$

$$\mathbf{Z} = \mathbf{Y}' \mathbf{Y} \text{ onde}$$

$$\mathbf{Y} = D_c^{-1/2} (\mathbf{P}' - \mathbf{cr}')' D_r^{-1/2}$$

Inércia Total

$$\text{InT} = \text{In}_c(\mathbf{R}) = \text{tr}(\mathbf{S})$$

$$= \text{In}_r(\mathbf{C}) = \text{tr}(\mathbf{Z})$$

Os passos algébricos para resolver o problema do ajuste são similares aos desenvolvidos na seção 2.3.3 e recaem na decomposição em valores singulares de  $\mathbf{X}$  ou  $\mathbf{Y}$ , dados por

$$\mathbf{X} = \mathbf{U} D_\alpha \mathbf{V}'$$

$$\mathbf{Y} = \mathbf{V} D_\alpha \mathbf{U}'$$

onde  $\mathbf{V}$  e  $\mathbf{U}$  são as matrizes cujas colunas nos dão as base do subespaço ótimo de  $\mathbf{R}^J$  e  $\mathbf{R}^I$

respectivamente

Coordenadas Principais

$$\mathbf{F} = \mathbf{D}_r^{-1/2} \mathbf{X} \mathbf{V}$$

$$\mathbf{F} = \mathbf{D}_r^{1/2} \mathbf{U} \mathbf{D}_\alpha$$

$$\mathbf{G} = \mathbf{D}_c^{1/2} \mathbf{X}' \mathbf{U}$$

$$\mathbf{G} = \mathbf{D}_c^{1/2} \mathbf{V} \mathbf{D}_\alpha$$

Coordenadas Padronizadas

$$\Phi = \mathbf{F} \mathbf{D}_\alpha^{-1}$$

$$\Gamma = \mathbf{G} \mathbf{D}_\alpha^{-1}$$

Fórmulas de Transição

em coordenadas principais

$$\mathbf{F} = \mathbf{R} \mathbf{G} \mathbf{D}_\alpha^{-1}$$

$$\mathbf{G} = \mathbf{C} \mathbf{F} \mathbf{D}_\alpha^{-1}$$

em coordenadas padronizadas

$$\mathbf{F} = \mathbf{R} \Gamma$$

$$\mathbf{G} = \mathbf{C} \Phi$$

## A.2 ACS DA MATRIZ $N_1$

Tomemos agora a matriz cujas colunas foram agrupadas, ou seja, a matriz  $N_1$ . E, para ela, apresentaremos, como foi feito na seção anterior, alguns conceitos e resultados. Neste caso, a matriz de dados é dada por:

$$N_1 = \begin{bmatrix} n_{1\bullet}^{11} & \cdots & n_{1\bullet}^{1L} \\ \vdots & & \vdots \\ n_{1_i\bullet}^{11} & \cdots & n_{1_i\bullet}^{1L} \\ \vdots & & \vdots \\ n_{1\bullet}^{H1} & \cdots & n_{1\bullet}^{HL} \\ \vdots & & \vdots \\ n_{1_{ih}\bullet}^{H1} & \cdots & n_{1_{ih}\bullet}^{HL} \end{bmatrix}$$

e a matriz de correspondência pode, então, ser colocada como

$$P_1 = 1/nN_1.$$

O vetor de massas dos perfis linhas e centróide dos perfis colunas de grupos é

$$r_1 = \left( \frac{n_{1\bullet}^{1\bullet}}{n}, \dots, \frac{n_{1_i\bullet}^{1\bullet}}{n}, \dots, \frac{n_{1\bullet}^{H\bullet}}{n}, \dots, \frac{n_{1_{ih}\bullet}^{H\bullet}}{n} \right),$$

O vetor de massas dos perfis colunas de grupos e centróide dos perfis linhas serão dados por

$$c_1 = \left( \frac{n_{\bullet\bullet}^{1\bullet}}{n}, \dots, \frac{n_{\bullet\bullet}^{L\bullet}}{n} \right),$$

enquanto o  $i$ -ésimo perfil linha do  $h$ -ésimo grupo será

$$(R_i^h)_1 = \left( \frac{n_{i\bullet}^{h1}}{n_{i\bullet}^{h\bullet}}, \dots, \frac{n_{i\bullet}^{hL}}{n_{i\bullet}^{h\bullet}} \right), \quad i = 1, \dots, I_h \text{ e } h = 1, \dots, H.$$

Colocaremos a Matriz dos perfis linhas como sendo

$$R_1 = [(R_1^1)_1' \dots (R_{1_i}^1)_1' \dots (R_1^H)_1' \dots (R_{1_{ih}}^H)_1']' = D_{r_1}^{-1} P_{1..},$$

o  $\ell$ -ésimo perfil coluna de grupo como

$$(C_\bullet^\ell)_1 = \left( \frac{n_{1\bullet}^{1\ell}}{n_{\bullet\bullet}^{1\ell}}, \dots, \frac{n_{1_i\bullet}^{1\ell}}{n_{\bullet\bullet}^{1\ell}}, \dots, \frac{n_{1\bullet}^{H1}}{n_{\bullet\bullet}^{1\ell}}, \dots, \frac{n_{1_{ih}\bullet}^{H\ell}}{n_{\bullet\bullet}^{1\ell}} \right), \quad \ell = 1, \dots, L.$$

e a matriz dos perfis colunas de grupo será dada por

$$C_1 = [(C_\bullet^1)_1' \dots (C_\bullet^L)_1']' = D_{c_1}^{-1} P_{1..}'.$$

A distância definida para a nuvem dos pontos-linhas no espaço  $R^J$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_{c_2}$  onde na diagonal temos os elementos do centróide, é dada por

$$d^2((R_i^h)_1, (R_i^{h'})_1) = \sum_{\ell=1}^L \sum_{j=1}^{J_\ell} \left( \frac{n_{ij}^{h\ell}}{n_{i\bullet}^{h\bullet}} - \frac{n_{ij}^{h'\ell}}{n_{i\bullet}^{h'\bullet}} \right)^2 \frac{n}{n_{i\bullet}^{h\bullet} n_{i\bullet}^{h'\bullet}}$$

$$= ((R_i^h)_1 - (R_i^{h'})_1)' D_{c_i}^{-1} ((R_i^h)_1 - (R_i^{h'})_1)$$

e distância definida para a nuvem dos pontos-colunas no espaço  $R^L$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_{r_i}$  onde na diagonal temos os elementos do centróide, é dada por

$$d^2((C_i^\ell)_1, (C_i^{\ell'})_1) = \sum_{h=1}^H \left( \frac{n_{i\bullet}^{h\ell}}{n_{i\bullet}^{h\bullet}} - \frac{n_{i\bullet}^{h\ell'}}{n_{i\bullet}^{h'\bullet}} \right)^2 \frac{n}{n_{i\bullet}^{h\bullet} n_{i\bullet}^{h'\bullet}}$$

$$= ((C_i^\ell)_1 - (C_i^{\ell'})_1)' D_{r_i}^{-1} ((C_i^\ell)_1 - (C_i^{\ell'})_1)$$

A Matriz que define o ajuste da nuvem do espaço  $R^L$  é dada por

$$S_1 = (P_1 - r_1 c_1)' D_{r_1}^{-1} (P_1 - r_1 c_1)' D_{c_1}^{-1}$$

$$S_1 = X_1' X_1 \text{ onde}$$

$$X_1 = D_{r_1}^{-1/2} (P_1 - r_1 c_1)' D_{c_1}^{-1/2}$$

e a Matriz que define o ajuste da nuvem do espaço  $R^L$  é dada por

$$Z_1 = (P_1' - c_1 r_1)' D_{c_1}^{-1} (P_1' - c_1 r_1)' D_{r_1}^{-1}$$

$$Z_1 = Y_1' Y_1 \text{ onde}$$

$$Y_1 = D_{c_1}^{-1/2} (P_1' - c_1 r_1)' D_{r_1}^{-1/2}$$

Inércia Total

$$\text{InT} = \text{In}_{c_1} (R_1) = \text{tr}(S_1)$$

$$= \text{In}_{r_1} (C_1) = \text{tr}(Z_1)$$

Os passos algébricos para resolver o problema do ajuste são similares aos desenvolvidos na seção 2.3.3 e recaem na decomposição em valores singulares de  $X_1$  ou  $Y_1$ , dados por

$$X_1 = U_1 D_{\eta_1} V_1'$$

$$Y_1 = V_1 D_{\eta_1} U_1'$$

onde  $V_1$  e  $U_1$  são as matrizes cujas colunas nos dão as base do subespaço ótimo de  $R^L$  e  $R^L$  respectivamente

## Coordenadas Principais

$$\mathbf{F}_1 = \mathbf{D}_{\eta_1}^{-1} \mathbf{X}_1 \mathbf{V}_1$$

$$\mathbf{F}_1 = \mathbf{D}_{\eta_1}^1 \mathbf{U}_1 \mathbf{D}_{\eta_1}$$

$$\mathbf{G}_1 = \mathbf{D}_{\epsilon_2}^1 \mathbf{X}_1' \mathbf{U}_1$$

$$\mathbf{G}_1 = \mathbf{D}_{\epsilon_2}^1 \mathbf{V}_1 \mathbf{D}_{\eta_1}$$

## Coordenadas Padronizadas

$$\Phi_1 = \mathbf{F}_1 \mathbf{D}_{\eta_1}^{-1}$$

$$\Gamma_1 = \mathbf{G}_1 \mathbf{D}_{\eta_1}^{-1}$$

## Fórmulas de Transição

em coordenadas principais

$$\mathbf{F}_1 = \mathbf{R}_1 \mathbf{G}_1 \mathbf{D}_{\eta_1}^{-1}$$

$$\mathbf{G}_1 = \mathbf{C}_1 \mathbf{F}_1 \mathbf{D}_{\eta_1}^{-1}$$

em coordenadas padronizadas

$$\mathbf{F}_1 = \mathbf{R}_1 \Phi_1$$

$$\mathbf{G}_1 = \mathbf{C}_1 \Gamma_1$$

### A.3 ACS DA MATRIZ $N_2$

Tomemos agora a matriz oriunda do agrupamento das linhas, ou seja, tomemos a matriz que chamamos de matriz  $N_2$ .

Neste caso, a matriz de dados é colocada como

$$N_2 = \begin{bmatrix} n_{\bullet 1}^{11} & \cdots & n_{\bullet J_1}^{11} & n_{\bullet 1}^{1L} & \cdots & n_{\bullet J_L}^{1L} \\ \vdots & & \vdots & \vdots & & \vdots \\ n_{\bullet 1}^{H1} & \cdots & n_{\bullet J_1}^{H1} & n_{\bullet 1}^{HL} & \cdots & n_{\bullet J_L}^{HL} \end{bmatrix}$$

e a matriz de correspondência é

$$P_2 = 1/nN_2.$$

O vetor de massas dos perfis linhas de grupos e centróide dos perfis colunas será denotado por

$$r_2 = \left( \frac{n_{\bullet\bullet}^{1\bullet}}{n}, \dots, \frac{n_{\bullet\bullet}^{H\bullet}}{n} \right),$$

enquanto o vetor de massas dos perfis colunas e centróide dos perfis linhas de grupos será

$$c_2 = \left( \frac{n_{\bullet 1}^{1\bullet}}{n}, \dots, \frac{n_{\bullet J_1}^{1\bullet}}{n}, \dots, \frac{n_{\bullet 1}^{L\bullet}}{n}, \dots, \frac{n_{\bullet J_L}^{L\bullet}}{n} \right).$$

Continuando, temos que o  $h$ -ésimo perfil linha de grupo é dado por

$$(R_{\bullet}^h)_2 = \left( \frac{n_{\bullet 1}^{h1}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet J_1}^{h1}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet 1}^{hL}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet J_L}^{hL}}{n_{\bullet\bullet}^{h\bullet}} \right), \quad h = 1, \dots, H$$

e a matriz dos perfis linhas de grupo por

$$R_2 = [(R_{\bullet}^1)_2', \dots, (R_{\bullet}^H)_2']' = D_{r_2}^{-1} P_2.$$

O  $j$ -ésimo perfil coluna do  $\ell$ -ésimo grupo é

$$(C_j^\ell)_2 = \left( \frac{n_{\bullet j}^{1\ell}}{n_{\bullet\bullet}^{1\ell}}, \dots, \frac{n_{\bullet j}^{H\ell}}{n_{\bullet\bullet}^{H\ell}} \right), \quad j = 1, \dots, J_\ell \text{ e } \ell = 1, \dots, L$$

e a matriz dos perfis colunas é dada por

$$C_2 = [(C_1^1)_2', \dots, (C_{J_1}^1)_2', \dots, (C_1^L)_2', \dots, (C_{J_L}^L)_2']' = D_{c_2}^{-1} P_2'.$$

A distância definida para a nuvem dos pontos-linhas no espaço  $R^J$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_{c_2}$  onde na diagonal temos os elementos do centróide, é dada por

$$\begin{aligned} d^2((R_{\bullet}^h)_2, (R_{\bullet}^{h'})_2) &= \sum_{\ell=1}^L \sum_{j=1}^{J_\ell} \left( \frac{n_{\bullet j}^{h\ell}}{n_{\bullet\bullet}^{h\ell}} - \frac{n_{\bullet j}^{h'\ell}}{n_{\bullet\bullet}^{h'\ell}} \right)^2 \frac{n}{n_{\bullet j}^{h\ell}} \\ &= ((R_{\bullet}^h)_2 - (R_{\bullet}^{h'})_2)' D_{c_2}^{-1} ((R_{\bullet}^h)_2 - (R_{\bullet}^{h'})_2) \end{aligned}$$

e distância definida para a nuvem dos pontos-colunas no espaço  $\mathbf{R}^H$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_{r_2}$  onde na diagonal temos os elementos do centróide. Esta distância é dada por

$$d^2((C_J^l)_2, (C_{j'}^{l'})_2) = \sum_{h=1}^H \left( \frac{n_{\bullet j}^{hl}}{n_{\bullet j}} - \frac{n_{\bullet j}^{h'l'}}{n_{\bullet j}} \right)^2 \frac{n}{n_{\bullet\bullet}^h}$$

$$= ((C_J^l)_2 - (C_{j'}^{l'})_2)' D_{r_2}^{-1} ((C_J^l)_2 - (C_{j'}^{l'})_2)$$

A Matriz que define o ajuste da nuvem do espaço  $\mathbf{R}^J$  é dada por

$$S_2 = (\mathbf{P}_2 - r_2 \mathbf{c}_2') D_{r_2}^{-1} (\mathbf{P}_2 - r_2 \mathbf{c}_2')' D_{c_2}^{-1}$$

$$S_2 = \mathbf{X}_2' \mathbf{X}_2 \text{ onde}$$

$$\mathbf{X}_2 = D_{r_2}^{-1/2} (\mathbf{P}_2 - r_2 \mathbf{c}_2') D_{c_2}^{-1/2}$$

e a Matriz que define o ajuste da nuvem do espaço  $\mathbf{R}^H$  é dada por

$$Z_2 = (\mathbf{P}_2' - c_2 r_2')' D_{c_2}^{-1} (\mathbf{P}_2' - c_2 r_2') D_{r_2}^{-1}$$

$$Z_2 = \mathbf{Y}_2' \mathbf{Y}_2 \text{ onde}$$

$$\mathbf{Y}_2 = D_{c_2}^{-1/2} (\mathbf{P}_2' - c_2 r_2')' D_{r_2}^{-1/2}$$

Inércia Total

$$\text{InT} = \text{In}_{c_2}(\mathbf{R}_2) = \text{tr}(S_2)$$

$$= \text{In}_{r_2}(\mathbf{C}_2) = \text{tr}(Z_2)$$

Os passos algébricos para resolver o problema do ajuste são similares aos desenvolvidos na seção 2.3.3 e recaem na decomposição em valores singulares de  $\mathbf{X}_2$  ou  $\mathbf{Y}_2$ , dados por

$$\mathbf{X}_2 = \mathbf{U}_2 D_{\eta_2} \mathbf{V}_2'$$

$$\mathbf{Y}_2 = \mathbf{V}_2 D_{\eta_2} \mathbf{U}_2'$$

onde  $\mathbf{V}_2$  e  $\mathbf{U}_2$  são as matrizes cujas colunas nos dão as base do subespaço ótimo de  $\mathbf{R}^J$  e  $\mathbf{R}^H$ , respectivamente.

Coordenadas Principais

$$\mathbf{F}_2 = D_{r_2}^{-1/2} \mathbf{X}_2 \mathbf{V}_2$$

$$\mathbf{F}_2 = D_{r_2}^{1/2} \mathbf{U}_2 D_{\eta_2}$$

$$\mathbf{G}_2 = \mathbf{D}_{c_2}^{1,2} \mathbf{X}_2' \mathbf{U}_2$$

$$\mathbf{G}_2 = \mathbf{D}_{c_2}^{1,2} \mathbf{V}_2 \mathbf{D}_{\eta_2}$$

Coordenadas Padronizadas

$$\Phi_2 = \mathbf{F}_2 \mathbf{D}_{\eta_2}^{-1}$$

$$\Gamma_2 = \mathbf{G}_2 \mathbf{D}_{\eta_2}^{-1}$$

Fórmulas de Transição

em coordenadas principais

$$\mathbf{F}_2 = \mathbf{R}_2 \mathbf{G}_2 \mathbf{D}_{\eta_2}^{-1}$$

$$\mathbf{G}_2 = \mathbf{C}_2 \mathbf{F}_2 \mathbf{D}_{\eta_2}^{-1}$$

em coordenadas padronizadas

$$\mathbf{F}_2 = \mathbf{R}_2 \Phi_2$$

$$\mathbf{G}_2 = \mathbf{C}_2 \Gamma_2$$

#### A.4 ACS DA MATRIZ $N_3$

Finalmente, ao tomarmos agora a matriz  $N_3$  cujas linhas e colunas estão agrupadas, desenvolvermos a ACS.

Notemos que, neste caso, a matriz de dados é dada por

$$N_3 = \begin{bmatrix} n_{\bullet\bullet}^{11} & \dots & n_{\bullet\bullet}^{1L} \\ \vdots & & \vdots \\ n_{\bullet\bullet}^{H1} & \dots & n_{\bullet\bullet}^{HL} \end{bmatrix}$$

e a matriz de Correspondência é

$$P_3 = 1/nN_3.$$

O vetor de massas dos perfis linha de grupos e centróide dos perfis coluna de grupos é

$$r_3 = \left( \frac{n_{\bullet\bullet}^{1\bullet}}{n}, \dots, \frac{n_{\bullet\bullet}^{H\bullet}}{n} \right),$$

O vetor de massas dos perfis coluna de grupos e centroide dos perfis linha de grupos é

$$c_3 = \left( \frac{n_{\bullet\bullet}^{\bullet 1}}{n}, \dots, \frac{n_{\bullet\bullet}^{\bullet L}}{n} \right),$$

O  $h$ -ésimo perfil linha de grupo, para a matriz  $N_3$ , é

$$(R_{\bullet}^h)_3 = \left( \frac{n_{\bullet\bullet}^{h1}}{n_{\bullet\bullet}^{h\bullet}}, \dots, \frac{n_{\bullet\bullet}^{hL}}{n_{\bullet\bullet}^{h\bullet}} \right), \quad h = 1, \dots, H$$

e a matriz dos perfis linha de grupo

$$R_3 = [(R_{\bullet}^1)_3, \dots, (R_{\bullet}^H)_3] = D_{r_3}^{-1} P_3.$$

Denotaremos o  $\ell$ -ésimo perfil coluna de grupo por

$$(C_{\bullet}^{\ell})_3 = \left( \frac{n_{\bullet\bullet}^{\bullet \ell}}{n_{\bullet\bullet}^{\bullet \ell}}, \dots, \frac{n_{\bullet\bullet}^{H\ell}}{n_{\bullet\bullet}^{\bullet \ell}} \right), \quad \ell = 1, \dots, L$$

enquanto a matriz dos perfis coluna de grupo será colocada como

$$C_3 = [(C_{\bullet}^1)_3, \dots, (C_{\bullet}^L)_3] = D_{c_3}^{-1} P_3.$$

A distância definida para a nuvem dos pontos-linhas no espaço  $R^L$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_{c_3}$ , onde na diagonal temos os elementos do centroide. Esta distância é dada por

$$\begin{aligned} d^2((R_{\bullet}^h)_3, (R_{\bullet}^{h'})_3) &= \sum_{\ell=1}^L \left( \frac{n_{\bullet\bullet}^{h\ell}}{n_{\bullet\bullet}^{h\bullet}} - \frac{n_{\bullet\bullet}^{h'\ell}}{n_{\bullet\bullet}^{h'\bullet}} \right)^2 \frac{n}{n_{\bullet\bullet}^{\bullet \ell}} \\ &= ((R_{\bullet}^h)_3 - (R_{\bullet}^{h'})_3)' D_{c_3}^{-1} ((R_{\bullet}^h)_3 - (R_{\bullet}^{h'})_3) \end{aligned}$$

e distância definida para a nuvem dos pontos-colunas no espaço  $\mathbf{R}^H$  é dada pela distância euclidiana generalizada, definida pela matriz diagonal  $D_{r_3}$ , onde na diagonal temos os elementos do centróide, que é dada por

$$\begin{aligned} d^2((C_{\bullet}^{\ell})_3, (C_{\bullet}^{\ell'})_3) &= \sum_{h=1}^H \left( \frac{n_{\bullet\bullet}^{h\ell}}{n_{\bullet\bullet}^{\ell}} - \frac{n_{\bullet\bullet}^{h\ell'}}{n_{\bullet\bullet}^{\ell'}} \right)^2 \frac{n}{n_{\bullet\bullet}^{h\ell}} \\ &= ((C_{\bullet}^{\ell})_3 - (C_{\bullet}^{\ell'})_3)' D_{r_3}^{-1} ((C_{\bullet}^{\ell})_3 - (C_{\bullet}^{\ell'})_3) . \end{aligned}$$

A Matriz que define o ajuste da nuvem do espaço  $\mathbf{R}^L$  é dada por

$$\mathbf{S}_3 = (\mathbf{P}_3 - \mathbf{r}_3 \mathbf{c}_3') D_{r_3}^{-1} (\mathbf{P}_3 - \mathbf{r}_3 \mathbf{c}_3')' D_{c_3}^{-1}$$

$$\mathbf{S}_3 = \mathbf{X}_3' \mathbf{X}_3 \text{ onde}$$

$$\mathbf{X}_3 = D_{r_3}^{-1/2} (\mathbf{P}_3 - \mathbf{r}_3 \mathbf{c}_3') D_{c_3}^{-1/2}$$

e a Matriz que define o ajuste da nuvem do espaço  $\mathbf{R}^H$  é dada por

$$\mathbf{Z}_3 = (\mathbf{P}_3' - \mathbf{c}_3 \mathbf{r}_3')' D_{c_3}^{-1} (\mathbf{P}_3' - \mathbf{c}_3 \mathbf{r}_3') D_{r_3}^{-1}$$

$$\mathbf{Z}_3 = \mathbf{Y}_3' \mathbf{Y}_3 \text{ onde}$$

$$\mathbf{Y}_3 = D_{c_3}^{-1/2} (\mathbf{P}_3' - \mathbf{c}_3 \mathbf{r}_3')' D_{r_3}^{-1/2}$$

Inércia Total

$$\text{InT} = \text{In}_{c_3}(\mathbf{R}_3) = \text{tr}(\mathbf{S}_3)$$

$$= \text{In}_{r_3}(\mathbf{C}_3) = \text{tr}(\mathbf{Z}_3)$$

Os passos algébricos para resolver o problema do ajuste são similares aos desenvolvidos na seção 2.3.3 e recaem na decomposição em valores singulares de  $\mathbf{X}_3$  ou  $\mathbf{Y}_3$ , dados por

$$\mathbf{X}_3 = \mathbf{U}_3 D_{\eta_3} \mathbf{V}_3'$$

$$\mathbf{Y}_3 = \mathbf{V}_3 D_{\eta_3} \mathbf{U}_3'$$

onde  $\mathbf{V}_3$  e  $\mathbf{U}_3$  são as matrizes cujas colunas nos dão as base do subespaço ótimo de  $\mathbf{R}^L$  e  $\mathbf{R}^H$ , respectivamente.

Coordenadas Principais

$$\mathbf{F}_3 = D_{r_3}^{-1/2} \mathbf{X}_3 \mathbf{V}_3$$

$$\mathbf{F}_3 = D_{r_3}^{1/2} \mathbf{U}_3 D_{\eta_3}$$

$$\mathbf{G}_3 = \mathbf{D}_{e_3}^{1,2} \mathbf{X}_3 \mathbf{U}_3$$

$$\mathbf{G}_3 = \mathbf{D}_{e_3}^{1,2} \mathbf{V}_3 \mathbf{D}_{\eta_3}$$

Coordenadas Padronizadas

$$\mathbf{\Phi}_3 = \mathbf{F}_3 \mathbf{D}_{\eta_3}^{-1}$$

$$\mathbf{\Gamma}_3 = \mathbf{G}_3 \mathbf{D}_{\eta_3}^{-1}$$

Fórmulas de Transição

em coordenadas principais

$$\mathbf{F}_3 = \mathbf{R}_3 \mathbf{G}_3 \mathbf{D}_{\eta_3}^{-1}$$

$$\mathbf{G}_3 = \mathbf{C}_3 \mathbf{F}_3 \mathbf{D}_{\eta_3}^{-1}$$

em coordenadas padronizadas

$$\mathbf{F}_3 = \mathbf{R}_3 \mathbf{\Phi}_3$$

$$\mathbf{G}_3 = \mathbf{C}_3 \mathbf{\Gamma}_3$$

## APÊNDICE B

## B.1 DESCRIÇÃO PARA O ALGORITMO DA AC

A Análise de Correspondência (AC) é uma técnica multivariada para análise exploratória de dados categorizados. Ela converte uma matriz de dados não negativos em um particular tipo de gráfico que exhibe as linhas e colunas da matriz como pontos de um espaço vetorial de dimensão menor que a original, de maneira que as associações entre as linhas, entre as colunas e entre linhas e colunas possam ser interpretadas. Podemos considerar a AC como sendo um algoritmo de redução de dados que fornece imagens simplificadas da realidade multidimensional e busca a melhor representação simultânea dos conjuntos de categorias da matriz de dados. A seguir descrevemos as expressões que nos possibilita montar o algoritmo da AC.

### I - PROCEDIMENTO PARA DETERMINAÇÃO DOS EIXOS

#### 1 - ENTRADA DOS DADOS

**N**                      **dim(IxJ)**

**I:** nº de linhas

**J:** nº de colunas

**N** é qualquer matriz de dados não negativos de dimensão IxJ, sendo que I é o número de linhas e J é o número de colunas. Os casos mais comuns são: contagens relativas ao cruzamento de duas multinomiais (ou tabela de contingência), ou as frequências relativas a uma multinomial em grupos representados por uma das entradas.  $N=[n_{ij}]$ .

#### 2 - DEFINIÇÃO DE VETORES AUXILIARES

**HI** = (1, ... , 1)<sup>T</sup>      **dim(Ix1)**

**HI** é um vetor de dimensão Ix1 com todos os elementos iguais a 1 (um).(1<sub>i</sub>)

**HJ** = (1, ... , 1)<sup>T</sup>      **dim(Jx1)**

**HJ** é um vetor de dimensão Jx1 com todos os elementos iguais a 1 (um).(1<sub>j</sub>)

**HK** = (1, ... , 1)<sup>T</sup>      **dim(Kx1)**

**HK** é um vetor de dimensão Kx1 com todos os elementos iguais a 1 (um).(1<sub>k</sub>)

#### 3 - CÁLCULO DOS DE ELEMENTOS MARGINAIS E TOTAL DA MATRIZ DE DADOS

**NT** = (HI)<sup>T</sup>\*N\*HJ    **escalar**

**NT** é o número de elementos da matriz de dados. (n)

$$\mathbf{CM} = \mathbf{N} * \mathbf{HJ} \quad \mathbf{dim(Ix1)}$$

**CM** é o vetor de dimensão  $I \times 1$  que nos dá a coluna marginal da matriz de dados, cada elemento desse vetor é a soma das colunas em cada linha. ( $n_i$ )

$$\mathbf{LM} = (\mathbf{HI})^T * \mathbf{N} \quad \mathbf{dim(Jx1)}$$

**LM** é o vetor de dimensão  $J \times 1$  que nos dá a linha marginal da matriz de dados, onde cada elemento desse vetor é a soma das linha em cada colunas. ( $n_j$ )

#### 4 - CÁLCULO DA MATRIZ DE CORRESPONDÊNCIA

$$\mathbf{P} = (1/\mathbf{NT}) * \mathbf{N} \quad \mathbf{dim(IxJ)}$$

**P** é uma matriz de dimensão  $I \times J$  com os elementos  $p_{ij} = n_{ij}/n$ . Quando **N** é uma tabela de contingência, **P** é uma matriz de frequências relativas.

#### 5 - CÁLCULO DO VETOR DE MASSAS DAS LINHAS

$$\mathbf{R} = \mathbf{P} * \mathbf{HJ} \quad \mathbf{dim(Ix1)}$$

**R** é um vetor de dimensão  $I \times 1$  com os elementos  $r_i = n_i/n$ . **R** na análise da linhas é o vetor de massas dos perfis linha e na análise das colunas ele é o centróide da nuvem de perfis coluna. É o vetor  $\mathbf{r} = (r_1, \dots, r_I)$ .

#### 6 - CÁLCULO DO VETOR DE MASSAS DAS COLUNAS

$$\mathbf{C} = \mathbf{P}^T * \mathbf{HI} \quad \mathbf{dim(Jx1)}$$

**C** é um vetor de dimensão  $J \times 1$  com os elementos  $c_j = n_j/n$ . **C** na análise da linhas é o centróide da nuvem de perfis linha e na análise das colunas ele é o vetor de massas dos perfis coluna. É o vetor  $\mathbf{c} = (c_1, \dots, c_J)$ .

#### 7 - CÁLCULO DA MATRIZ DOS VALORES ESPERADOS

$$\mathbf{E} = \mathbf{NT} * \mathbf{R} * \mathbf{C}^T \quad \mathbf{dim(IxJ)}$$

**E**=[**E1**...**EJ**] é uma matriz de dimensão  $I \times J$  com os elementos  $e_{ij} = (n_i n_j)/n$ . Seus elementos são os valores esperados segundo a distribuição Qui-quadrado.

No SAS é o valor esperado da estatística qui-quadrado (Chi-Square Statistic Expected Values )

## 8 - CÁLCULO DA MATRIZ DOS VALORES OBSERVADOS MENOS OS VALORES ESPERADOS

$$OE = N - E \quad \text{dim}(I \times J)$$

**O** é uma matriz de dimensão  $I \times J$  com os elementos  $oe_{ij} = n - (n_i \cdot n_j) / n$ . Seus elementos são os valores observados da matriz **N**, menos os valores esperados da matriz **E**. No SAS é o valor observado menos valor esperado (Observed Minus Expected Values)

$$QOE = [QOE1 | \dots | QOEJ]$$

**QOE** é a matriz onde cada elemento é o quadrado do correspondente elemento de **OE**.

## 9 - CÁLCULO DA MATRIZ DAS CONTRIBUIÇÕES TOTAIS DO QUI-QUADRADO

$$Q1 = \text{inv}(\text{diag}(E1)) * QOE1$$

**Q1** é o vetor de dimensão  $I \times 1$  onde cada elemento é o quociente entre o correspondente elemento da primeira coluna de **QOE** e os da primeira coluna de **E1**. No SAS é a contribuição total à estatística qui-quadrado (Contributions to the Total Chi-Square Statistic)

$$QJ = \text{inv}(\text{diag}(EJ)) * QOEJ$$

**QK** é o vetor de dimensão  $I \times 1$  onde cada elemento é o quociente entre o correspondente elemento da  $j$ -ésima coluna de **QOE** e os da  $j$ -ésima coluna de **E1**.

$$Q = [Q1 | \dots | QJ]$$

**Q** é a matriz de dimensão  $I \times J$  onde cada elemento é o quociente entre o correspondente elemento de **QOE** e os de **E1**.

## 10 - DEFINIÇÃO DE MATRIZES DIAGONAIS

$$DR = \text{diag}(R) \quad \text{dim}(I \times I)$$

**DR** é uma matriz diagonal de dimensão  $I \times I$  com os elementos da diagonal iguais aos elementos do vetor marginal **R**.  $D_r$

$$DIR = \text{inv}(DR) \quad \text{dim}(I \times I)$$

**DIR** é a matriz diagonal de dimensão  $I \times I$  com os elementos da diagonal iguais ao inverso dos elementos do vetor marginal **R**.  $D_r^{-1}$

$$\mathbf{DRIR} = \text{sqrt}(\mathbf{DIR}) \quad \text{dim}(\mathbf{I} \times \mathbf{I})$$

$\mathbf{DRIR}$  é a matriz diagonal de dimensão  $\mathbf{I} \times \mathbf{I}$  com os elementos da diagonal iguais à raiz quadrada do inverso dos elementos do vetor marginal  $\mathbf{R}$ .  $D_r^{-1/2}$

$$\mathbf{DC} = \text{diag}(\mathbf{C}) \quad \text{dim}(\mathbf{J} \times \mathbf{J})$$

$\mathbf{DC}$  é uma matriz diagonal de dimensão  $\mathbf{J} \times \mathbf{J}$  com os elementos da diagonal iguais aos elementos do vetor marginal  $\mathbf{C}$ .  $D_c$

$$\mathbf{DIC} = \text{inv}(\mathbf{DC}) \quad \text{dim}(\mathbf{J} \times \mathbf{J})$$

$\mathbf{DIC}$  é a matriz diagonal de dimensão  $\mathbf{J} \times \mathbf{J}$  com os elementos da diagonal iguais ao inverso dos elementos do vetor marginal  $\mathbf{C}$ .  $D_c^{-1}$

$$\mathbf{DRIC} = \text{sqrt}(\mathbf{DIC}) \quad \text{dim}(\mathbf{J} \times \mathbf{J})$$

$\mathbf{DRIC}$  é a matriz diagonal de dimensão  $\mathbf{J} \times \mathbf{J}$  com os elementos da diagonal iguais à raiz quadrada do inverso dos elementos do vetor marginal  $\mathbf{C}$ .  $D_c^{-1/2}$

## 11 - CÁLCULO DA MATRIZ DOS PERFIS LINHA

$$\mathbf{RR} = \mathbf{DIR} * \mathbf{P} \quad \text{dim}(\mathbf{I} \times \mathbf{J})$$

$\mathbf{RR}$  é uma matriz dos perfis linha de dimensão  $\mathbf{I} \times \mathbf{J}$ .  $\mathbf{R} = D_r^{-1} \mathbf{P}$ . No SAS é o perfis das linhas (*Row Profiles*)

## 12 - CÁLCULO DA MATRIZ DOS PERFIS COLUNA

$$\mathbf{CC} = \mathbf{DIC} * \mathbf{P}^T \quad \text{dim}(\mathbf{I} \times \mathbf{J})$$

$\mathbf{CC}$  é uma matriz dos perfis coluna de dimensão  $\mathbf{I} \times \mathbf{J}$ .  $\mathbf{C} = D_c^{-1} \mathbf{P}$ . No SAS é o perfis das colunas (*Column Profiles*)

## 13 - CÁLCULO DA MATRIZ DE CENTRALIZAÇÃO, PONDERADA E TRANSFORMADA

$$\mathbf{X} = \mathbf{DRIR} * [\mathbf{P} - \mathbf{R} * \mathbf{C}^T] * \mathbf{DRIC} \quad \text{dim}(\mathbf{I} \times \mathbf{J})$$

$\mathbf{X}$  é a matriz de centralização dos perfis coluna onde foi feita uma transformação e a ponderação. Esta matriz de dimensão  $\mathbf{I} \times \mathbf{J}$ , é onde se aplica a sub-rotina da decomposição em valores singulares (SVD), transformando em  $\mathbf{X} = \mathbf{U} \mathbf{D}_x \mathbf{V}^T$ .  $\mathbf{X} = D_r^{-1/2} [\mathbf{P} - \mathbf{r} \mathbf{c}^T] D_c^{-1/2}$

**14 - DECOMPOSIÇÃO EM VALORES SINGULARES DA MATRIZ X**

$$[V,A,U] = \text{svd}(X,0)$$

$\text{svd}(X,0)$  é a decomposição em valores singulares de forma que são mostrados apenas os K valores singulares diferente de zero dado em A em ordem decrescente e os seus vetores associados das linhas V e das colunas U.

$$V = [ V1 \mid \dots \mid VK ] \quad \text{dim}(J \times K)$$

**K:** nº de valores singulares não nulos.

V é uma matriz de dimensão JxK onde nas colunas estão os autovetores das linhas associados aos K valores singulares diferentes de zeros. Estes vetores nos dão as direções de maior inércia, em ordem decrescente, da nuvem de perfis linha. Se as linhas e as colunas da matriz de dados forem linearmente independente o número de valores singulares não nulos coincide com o mínimo entre I e J menos um.  $K = \min\{I, J\} - 1$

$$V1 \quad \text{dim}(J \times 1)$$

V1 é o vetor de dimensão Jx1 que nos dá a direção do eixo de maior inércia da nuvem de perfis linha.  $v_1 = (v_{11}, \dots, v_{J1})$

$$VK \quad \text{dim}(J \times 1)$$

VK é o vetor de dimensão Jx1 que nos dá a direção do eixo de menor inércia (direção do K-ésimo eixo de maior inércia) da nuvem de perfis linha.  $v_k = (v_{1k}, \dots, v_{Jk})$

$$U = [ U1 \mid \dots \mid UK ] \quad \text{dim}(I \times K)$$

U é uma matriz de dimensão IxK onde nas colunas estão os autovetores das colunas associados aos K valores singulares diferentes de zeros. Estes vetores nos dão as direções de maiores inércias, em ordem decrescente, da nuvem de perfis coluna.

$$U1 \quad \text{dim}(I \times 1)$$

U1 é o vetor de dimensão Ix1 que nos dá a direção do eixo de maior inércia da nuvem de perfis coluna.  $u_1 = (u_{11}, \dots, u_{I1})$

$$UK \quad \text{dim}(I \times 1)$$

UK é o vetor de dimensão Ix1 que nos dá a direção do eixo de menor inércia (direção do K-ésimo eixo de maior inércia) da nuvem de perfis coluna.  $u_k = (u_{1k}, \dots, u_{Ik})$

$$A = (A_1, \dots, A_K) \quad \text{dim}(K \times 1)$$

$A$  é o vetor de dimensão  $K \times 1$  que nos dá os valores singulares da matriz  $X$  em ordem decrescente.  $A = (\alpha_1, \dots, \alpha_K)$  No SAS são os valores singulares (Singular Values) dados em inércia e decomposição da estatística qui-quadrado (Inertia and Chi-Square Decomposition)

$$DA = \text{diag}(A) \quad \text{dim}(K \times K)$$

$DA$  é uma matriz diagonal de dimensão  $K \times K$  onde na diagonal principal estão os valores singulares dados pelo vetor  $A$ .  $D_{\alpha}$

### 15 - CÁLCULO DAS COORDENADAS PRINCIPAIS DOS PERFIS LINHA

$$F = DRIR * U * DA \quad \text{dim}(I \times K)$$

$F = [F_1 | \dots | F_K]$  é uma matriz de dimensão  $I \times K$  onde nas colunas temos  $K$  vetores que nos dão as coordenadas principais dos  $I$  perfis linha em cada eixo.

$$F_1 \quad \text{dim}(I \times 1)$$

$F_1$  é o vetor de dimensão  $I \times 1$  que nos dá as coordenadas (ou projeções) dos  $I$  perfis linha no eixo de maior inércia da nuvem.  $F_1 = (f_{11}, \dots, f_{I1})$

$$F_K \quad \text{dim}(I \times 1)$$

$F_K$  é o vetor de dimensão  $I \times 1$  que nos dá as coordenadas (ou projeções) dos  $I$  perfis linha no eixo de menor inércia ( $K$ -ésimo eixo) da nuvem.  $F_K = (f_{1k}, \dots, f_{Ik})$

### 16 - CÁLCULO DE VETORES COM ELEMENTOS AO QUADRADO

$$DF_1 = \text{diag}(F_1) \quad \text{dim}(I \times I)$$

$$DQF_1 = DF_1 * DF_1 \quad \text{dim}(I \times I)$$

$$QF_1 = DQF_1 * HI \quad \text{dim}(I \times 1)$$

$QF_1$  é o vetor de dimensão  $I \times 1$  que nos dá o quadrado de cada componente do vetor de coordenadas  $F_1$ .  $F_1^2 = (f_{11}^2, \dots, f_{I1}^2)$

$$DF_K = \text{diag}(F_K) \quad \text{dim}(I \times I)$$

$$DQF_K = DF_K * DF_K \quad \text{dim}(I \times I)$$

$$QF_K = DQF_K * HI \quad \text{dim}(I \times 1)$$

$QF_K$  é o vetor de dimensão  $I \times 1$  que nos dá o quadrado de cada componente do vetor de coordenadas  $F_K$ .  $F_K^2 = (f_{1k}^2, \dots, f_{Ik}^2)$

$$\mathbf{DL} = \mathbf{DA} * \mathbf{DA} \quad \text{dim}(\mathbf{K} \times \mathbf{K})$$

$$\mathbf{L} = \mathbf{DL} * \mathbf{HK} \quad \text{dim}(\mathbf{I} \times \mathbf{I})$$

$\mathbf{L}$  é o vetor de dimensão  $\mathbf{K} \times \mathbf{I}$  que nos dá o quadrado de cada componente do vetor  $\mathbf{A}$  dos valores singulares.  $\mathbf{L} = (\mathbf{L}_1, \dots, \mathbf{L}_K)^T$  é o vetor cujas componentes são os autovalores. Estes autovalores são os quantificadores das componentes da inércia total da nuvem dos perfis linha, em cada eixo;  $\mathbf{L} = (\lambda_1, \dots, \lambda_K)'$ .

$$\mathbf{L}_1 \quad \text{escalar}$$

$\mathbf{L}_1$  é o escalar que nos dá a componente da inércia total da nuvem de perfis linha na direção do primeiro eixo. Esta componente é chamada de inércia principal do primeiro eixo;  $\lambda_1 = \alpha_1^2$ .

$$\mathbf{L}_K \quad \text{escalar}$$

$\mathbf{L}_K$  é o escalar que nos dá a inércia da nuvem de perfis linha na direção do  $\mathbf{K}$ -ésimo eixo. Esta componente é chamada de inércia principal do  $\mathbf{K}$ -ésimo eixo;  $\lambda_k = \alpha_k^2$ .

### 17 - CÁLCULO DA COMPONENTE DA INÉRCIA DE CADA PERFIL LINHA EM CADA DIREÇÃO

$$\mathbf{CIR}_1 = \mathbf{DR} * \mathbf{QF}_1 \quad \text{dim}(\mathbf{I} \times \mathbf{I})$$

$\mathbf{CIR}_1$  é o vetor de dimensão  $\mathbf{I} \times \mathbf{I}$  que nos dá a componente da inércia de todos os perfis na direção do primeiro eixo;  $[\text{In}_c(\mathbf{R})]_{v1} = D_r F_1^2$ .

$$\mathbf{CIR}_K = \mathbf{DR} * \mathbf{QF}_K \quad \text{dim}(\mathbf{I} \times \mathbf{I})$$

$\mathbf{CIR}_K$  é o vetor de dimensão  $\mathbf{I} \times \mathbf{I}$  que nos dá a componente da inércia de todos os perfis na direção do  $\mathbf{K}$ -ésimo eixo;  $[\text{In}_c(\mathbf{R})]_{vK} = ([\text{In}_c(\mathbf{R})]_{1k}, \dots, [\text{In}_c(\mathbf{R})]_{ik})' [\text{In}_c(\mathbf{R})]_{ik} = r_i F_{ik}^2$ .

### 18 - CÁLCULO DA INÉRCIA DOS PERFIS LINHA

$$\mathbf{IR} = \mathbf{DR} * \sum_k \mathbf{QF}_K \quad \text{dim}(\mathbf{I} \times \mathbf{I})$$

$\mathbf{IR}$  é o vetor de dimensão  $\mathbf{I} \times \mathbf{I}$  que nos dá a inércia de todos os perfis linha na nuvem.  $\text{In}_c(\mathbf{R}_i) = \sum_k [\text{In}_c(\mathbf{R}_i)]_{vk} = r_i \sum_k F_{ik}^2$ . No SAS as inércias dos pontos são dadas em termos percentuais, fazendo o quociente entre a inércia de cada ponto e a inércia total da nuvem. Dada no resumo estatístico para os pontos-linha (Summary Statistics for the Row Points).

$$\mathbf{DIR} = \text{diag}(\mathbf{IR}) \quad \text{dim}(\mathbf{I} \times \mathbf{I})$$

**19 - CÁLCULO DA CONTRIBUIÇÃO RELATIVA DOS PERFIS LINHA EM CADA EIXO**

$$\mathbf{CRR1} = \text{inv}(\mathbf{DIR}) * \mathbf{CIR1} \quad \text{dim}(\mathbf{Ix1})$$

**CRR1** é o vetor de dimensão  $\mathbf{Ix1}$  que nos dá a contribuição relativa dos perfis linhas no primeiro eixo. Cada componente deste vetor nos mostra o quanto cada perfil linha está contribuindo para a inércia do primeiro eixo comparada com a inércia de cada perfil em todo o espaço.  $\mathbf{cr}_1 = (cr_{11}, \dots, cr_{1n})$

$$cr_{i1} = [In_c(\mathbf{R}_i)]_1 / \sum_k [In_c(\mathbf{R}_i)]_k = [In_c(\mathbf{R}_i)]_1 / In_c(\mathbf{R}_i) = (r_i F_{i1}^2) / (\sum_k r_i F_{ik}^2).$$

$$\mathbf{CRRK} = \text{inv}(\mathbf{DIR}) * \mathbf{CIRK} \quad \text{dim}(\mathbf{Ix1})$$

**CRRK** é o vetor de dimensão  $\mathbf{Ix1}$  que nos dá a contribuição relativa dos perfis linhas no K-ésimo eixo. Cada componente deste vetor nos mostra o quanto cada perfil linhas contribui da sua inércia individual, para a composição da inércia principal deste eixo. Isso mostra o quão bem representado está cada ponto, neste eixo.  $\mathbf{cr}_k = (cr_{1k}, \dots, cr_{nk})$ .

$cr_{ik} = (r_i F_{ik}^2) / (\sum_k r_i F_{ik}^2) = F_{ik}^2 / F_i^2 = \cos^2 \theta$  onde  $\theta$  é o ângulo formado entre o eixo  $k$  e o vetor que aponta para o  $i$ -ésimo perfil linha.

**20 - CÁLCULO DA CONTRIBUIÇÃO ABSOLUTA DOS PERFIS LINHA EM CADA EIXO**

$$\mathbf{CAR1} = (1/L1) * \mathbf{CIR1} \quad \text{dim}(\mathbf{Ix1})$$

**CAR1** é o vetor de dimensão  $\mathbf{Ix1}$  que nos dá a contribuição absoluta dos perfis linhas no primeiro eixo. Cada componente deste vetor nos mostra o quanto as componentes de cada perfil linha neste eixo, representam na inércia principal do primeiro eixo. Isso mostra o quanto cada perfil linha contribui para a inércia principal do eixo.  $\mathbf{ca}_1 = (ca_{11}, \dots, ca_{1n})$   $ca_{11} = [In_c(\mathbf{R}_1)]_1 / \sum_i [In_c(\mathbf{R}_i)]_1 = [In_c(\mathbf{R}_1)]_1 / \lambda_1 = (r_1 F_{11}^2) / (\lambda_1)$ .

$$\mathbf{CARK} = (1/LK) * \mathbf{CIRK} \quad \text{dim}(\mathbf{Ix1})$$

**CARK** é o vetor de dimensão  $\mathbf{Ix1}$  que nos dá a contribuição absoluta dos perfis linhas no K-ésimo eixo. Cada componente deste vetor nos mostra o quanto as componentes de cada perfil linha neste eixo, representam na inércia principal do K-ésimo eixo. Isso mostra o quanto cada perfil linha contribui para a inércia principal do eixo.  $\mathbf{ca}_k = (ca_{1k}, \dots, ca_{nk})$   $ca_{1k} = [In_c(\mathbf{R}_1)]_k / \sum_i [In_c(\mathbf{R}_i)]_k = [In_c(\mathbf{R}_1)]_k / \lambda_k = (r_1 F_{1k}^2) / (\lambda_k)$ .

No SAS é dado em contribuição parcial à inércia dos pontos-linha (Partial Contributions to Inertia for the Row Points)

Aqui sugiro que na saída seja apresentado um vetor cujos elementos seja os nomes dados às linhas na ordem decrescente dos valores das contribuições absolutas. No SAS é apresentado os índices das coordenadas que mais contribuem com a inércia das pontos linha (Indices of the Coordinates that Contribute Most to Inertia for the Row Points)

Aparecendo o número da dimensão em que o ponto contribui muito e zero na que contribui pouco, se aparecer zero nas duas colunas é porque o ponto contribui pouco nestas duas dimensões, se aparecer o número da dimensão que mais contribui

nas duas colunas é porque o ponto contribui muito nas duas dimensões, mas contribui mais na especificada.

## 21 - CÁLCULO DAS COORDENADAS PADRONIZADAS DOS PERFIS LINHA

$$FP = F * \text{inv}(DA)$$

## 22 - CÁLCULO DAS COORDENADAS PRINCIPAIS DOS PERFIS COLUNA

$$G = \text{DRIC} * V * DA \quad \text{dim}(J \times K)$$

$G = [G1 | \dots | GK]$  é uma matriz de dimensão  $J \times K$  onde nas colunas temos  $K$  vetores que nos dão as coordenadas principais dos  $J$  perfis coluna em cada eixo.

$$G1 \quad \text{dim}(J \times 1)$$

$G1$  é o vetor de dimensão  $J \times 1$  que nos dá as coordenadas (ou projeções) dos  $J$  perfis coluna no eixo de maior inércia da nuvem.

$$GK \quad \text{dim}(J \times 1)$$

$GK$  é o vetor de dimensão  $J \times 1$  que nos dá as coordenadas (ou projeções) dos  $J$  perfis coluna no eixo de menor inércia ( $K$ -ésimo eixo) da nuvem.

## 23 - CÁLCULO DE VETORES COM ELEMENTOS AO QUADRADO

$$DG1 = \text{diag}(G1) \quad \text{dim}(J \times J)$$

$$DQG1 = DG1 * DG1 \quad \text{dim}(J \times J)$$

$$QG1 = DQG1 * HJ \quad \text{dim}(J \times 1)$$

$QG1$  é o vetor de dimensão  $J \times 1$  que nos dá o quadrado de cada componente do vetor de coordenadas  $G1$ .

$$DGK = \text{diag}(GK) \quad \text{dim}(J \times J)$$

$$DQGK = DGK * DGK \quad \text{dim}(J \times J)$$

$$QGK = DQGK * HJ \quad \text{dim}(J \times 1)$$

$QGK$  é o vetor de dimensão  $J \times 1$  que nos dá o quadrado de cada componente do vetor de coordenadas  $GK$ .

## 24 - CÁLCULO DA COMPONENTE DA INÉRCIA DE CADA PERFIL COLUNA EM CADA DIREÇÃO

$$CIC1 = DC * QG1 \quad \text{dim}(J \times 1)$$

**CIC1** é o vetor de dimensão  $J \times 1$  que nos dá a componente da inércia de todos os perfis coluna na direção do primeiro eixo.

$$\mathbf{CICK} = \mathbf{DC} * \mathbf{QGK} \quad \text{dim}(J \times 1)$$

**CICK** é o vetor de dimensão  $J \times 1$  que nos dá a componente da inércia de todos os perfis coluna na direção do K-ésimo eixo.

## 25 - CÁLCULO DA INÉRCIA DOS PERFIS COLUNA

$$\mathbf{IC} = \mathbf{DC} * \sum_k \mathbf{QGK} \quad \text{dim}(J \times 1)$$

**IC** é o vetor de dimensão  $J \times 1$  que nos dá a inércia de todos os perfis coluna na nuvem.

$$\mathbf{DIC} = \text{diag}(\mathbf{IC}) \quad \text{dim}(J \times J)$$

## 26 - CÁLCULO DA CONTRIBUIÇÃO RELATIVA DOS PERFIS COLUNA EM CADA EIXO

$$\mathbf{CRC1} = \text{inv}(\mathbf{DIC}) * \mathbf{CIC1} \quad \text{dim}(J \times 1)$$

**CRC1** é o vetor de dimensão  $J \times 1$  que nos dá a contribuição relativa dos perfis colunas no primeiro eixo. Cada componente deste vetor nos mostra o quanto cada perfil colunas está contribuindo para a inércia do primeiro eixo comparada com a inércia de cada perfil em todo o espaço.

$$\mathbf{CRCK} = \text{inv}(\mathbf{DIC}) * \mathbf{CICK} \quad \text{dim}(J \times 1)$$

**CRC1** é o vetor de dimensão  $J \times 1$  que nos dá a contribuição relativa dos perfis colunas no K-ésimo eixo. Cada componente deste vetor nos mostra o quanto cada perfil coluna contribui da sua inércia individual, para a composição da inércia principal deste eixo. Isso mostra o quão bem representado está cada ponto, neste eixo.

## 26 - CÁLCULO DA CONTRIBUIÇÃO ABSOLUTA DOS PERFIS COLUNA EM CADA EIXO

$$\mathbf{CAC1} = (1/L1) * \mathbf{CIC1} \quad \text{dim}(J \times 1)$$

**CAC1** é o vetor de dimensão  $J \times 1$  que nos dá a contribuição absoluta dos perfis colunas no primeiro eixo. Cada componente deste vetor nos mostra o quanto as componentes de cada perfil coluna neste eixo, representa na inércia principal do primeiro eixo. Isso mostra o quanto cada perfil coluna contribui para a inércia principal do eixo.

$$\mathbf{CACK} = (1/LK) * \mathbf{CICK} \quad \text{dim}(J \times 1)$$

**ACK** é o vetor de dimensão  $J \times 1$  que nos dá a contribuição absoluta dos perfis

coluna no K-ésimo eixo. Cada componente deste vetor nos mostra o quanto as componentes de cada perfil coluna neste eixo, representa na inércia principal do K-ésimo eixo. Isso mostra o quanto cada perfil coluna contribui para a inércia principal do eixo.

## 27 - CÁLCULO DAS COORDENADAS PADRONIZADAS DOS PERFIS COLUNA

$$GP = G * \text{inv}(DA) \quad \text{dim}(J \times K)$$

## II - PROCEDIMENTO PARA INCLUIR PONTOS-LINHAS SUPLEMENTARES

### 1 - ENTRADA DOS DADOS

$$NSR \quad \text{dim}(S \times J)$$

S: nº de linhas

J: nº de colunas

NSR é qualquer matriz de dados não negativos de dimensão  $S \times J$ , sendo que S é o número de linhas que esta sendo considerada como ponto suplementares e J é o número de colunas.  $NSR = [n_{sj}]$ .

### 2 - DEFINIÇÃO DE VETORES AUXILIARES

$$HS = (1, \dots, 1)^T \quad \text{dim}(S \times 1)$$

HS é um vetor de dimensão  $S \times 1$  com todos os elementos iguais a 1 (um). ( $1_s$ )

### 3 - CÁLCULO DOS DE ELEMENTOS MARGINAIS E TOTAL DA MATRIZ DE DADOS

$$NTRS = (HS)^T * N * HJ \quad \text{escalar}$$

NTRS é o número de elementos da matriz de dados. ( $n_s$ )

$$CMRS = NRS * HJ \quad \text{dim}(S \times 1)$$

CMRS é o vetor de dimensão  $1 \times 1$  que nos dá a coluna marginal da matriz de dados, cada elementos desse vetor é a soma das colunas em cada linha. ( $n_s$ )

$$LMRS = (HS)^T * N \quad \text{dim}(J \times 1)$$

LMRS é o vetor de dimensão  $J \times 1$  que nos dá a linha marginal da matriz de dados, onde cada elemento desse vetor é a soma das linha em cada colunas. ( $n_{sj}$ )

### 4 - CÁLCULO DA MATRIZ DE CORRESPONDÊNCIA

$$PRS = (1/NTRS) * NRS \quad \text{dim}(S \times J)$$

$P$  é uma matriz de dimensão  $I \times J$  com os elementos  $p_{ij} = n_{ij}/n$ . Quando  $N$  é uma tabela de contingência,  $P$  é uma matriz de frequências relativas.

## 5 - CÁLCULO DO VETOR DE MASSAS DAS LINHAS

$$RS = PRS * HJ \quad \text{dim}(S \times I)$$

$RS$  é um vetor de dimensão  $I \times 1$  com os elementos  $rs_i = n_{s_i}/n_s$ .  $rs = (rs_1, \dots, rs_I)$ .

## 10 - DEFINIÇÃO DE MATRIZES DIAGONAIS

$$DR = \text{diag}(RS) \quad \text{dim}(S \times I)$$

$DRS$  é uma matriz diagonal de dimensão  $I \times I$  com os elementos da diagonal iguais aos elementos do vetor marginal  $R$ .  $D_r$

$$DIRS = \text{inv}(DRS) \quad \text{dim}(S \times I)$$

$DIRS$  é a matriz diagonal de dimensão  $I \times I$  com os elementos da diagonal iguais ao inverso dos elementos do vetor marginal  $RS$ .  $D_{rs}^{-1}$

## 11 - CÁLCULO DA MATRIZ DOS PERFIS LINHA

$$RRS = DIRS * PRS \quad \text{dim}(S \times J)$$

$RR$  é uma matriz dos perfis linha de dimensão  $I \times J$ .  $RS = D_{rs}^{-1} PRS$

## 28 - CÁLCULO DOS COORDENADAS PRINCIPAIS DOS PERFIS LINHA SUPLEMENTARES

$$FS = RRS * GP \quad \text{dim}(S \times K)$$

$FS = [FS1 | \dots | FSK]$  é uma matriz de dimensão  $S \times K$  onde nas colunas temos  $K$  vetores que nos dão as coordenadas principais dos  $S$  perfis linha suplementares em cada eixo.

$$FS1 \quad \text{dim}(S \times 1)$$

$FS1$  é o vetor de dimensão  $S \times 1$  que nos dá as coordenadas (ou projeções) dos  $S$  perfis linha suplementares no eixo de maior inércia da nuvem.  $F_1 = (f_{11}, \dots, f_{I1})$

$$FSK \quad \text{dim}(S \times 1)$$

$FSK$  é o vetor de dimensão  $S \times 1$  que nos dá as coordenadas (ou projeções) dos  $S$  perfis linha suplementares no eixo de menor inércia ( $K$ -ésimo eixo) da nuvem.  $F_k = (f_{1k}, \dots, f_{Ik})$

**16 - CÁLCULO DE VETORES COM ELEMENTOS AO QUADRADO**

$$\mathbf{DFS1} = \text{diag}(\mathbf{FS1}) \quad \text{dim}(\mathbf{SxS})$$

$$\mathbf{DQFS1} = \mathbf{DFS1} * \mathbf{DFS1} \quad \text{dim}(\mathbf{SxS})$$

$$\mathbf{QFS1} = \mathbf{DQFS1} * \mathbf{HS} \quad \text{dim}(\mathbf{Sx1})$$

$\mathbf{QFS1}$  é o vetor de dimensão  $\mathbf{Sx1}$  que nos dá o quadrado de cada componente do vetor de coordenadas  $\mathbf{FS1}$ .  $\mathbf{FS1}^2 = (f_{s11}^2, \dots, f_{s1n}^2)$

$$\mathbf{DFSK} = \text{diag}(\mathbf{FSK}) \quad \text{dim}(\mathbf{SxS})$$

$$\mathbf{DQFSK} = \mathbf{DFSK} * \mathbf{DFSK} \quad \text{dim}(\mathbf{SxS})$$

$$\mathbf{QFSK} = \mathbf{DQFSK} * \mathbf{HS} \quad \text{dim}(\mathbf{Sx1})$$

$\mathbf{QFSK}$  é o vetor de dimensão  $\mathbf{Sx1}$  que nos dá o quadrado de cada componente do vetor de coordenadas  $\mathbf{FSK}$ .  $\mathbf{FSK}^2 = (f_{s1k}^2, \dots, f_{s1n}^2)$

$$\mathbf{SQFS} = \mathbf{QFS1} + \dots + \mathbf{QFSK} \quad \text{dim}(\mathbf{Sx1})$$

$$\mathbf{DSQFS} = \text{diag}(\mathbf{SQFS}) \quad \text{dim}(\mathbf{SxS})$$

**30 - CÁLCULO DA CONTRIBUIÇÃO RELATIVA DOS PERFIS LINHA SUPLEMENTARES EM CADA EIXO**

$$\mathbf{CRRS1} = \text{inv}(\mathbf{DSQFS}) * \mathbf{QFS1} \quad \text{dim}(\mathbf{Sx1})$$

$\mathbf{CRRS1}$  é o vetor de dimensão  $\mathbf{Sx1}$  que nos dá a contribuição relativa dos perfis linhas suplementares no primeiro eixo. Cada componente deste vetor nos mostra o quanto cada perfil linha suplementar está bem representado no primeiro eixo;  $\mathbf{cr}_1 = (cr_{11}, \dots, cr_{1n})$ .

$$cr_{1i} = [\ln_c(\mathbf{R}_i)]_1 / \sum_k [\ln_c(\mathbf{R}_i)]_k = [\ln_c(\mathbf{R}_i)]_1 / \ln_c(\mathbf{R}_i) = (r_i F_{i1}^2) / (\sum_k r_i F_{ik}^2)$$

$$\mathbf{CRRSK} = \text{inv}(\mathbf{DSQFS}) * \mathbf{QFSK} \quad \text{dim}(\mathbf{Ix1})$$

$\mathbf{CRRK}$  é o vetor de dimensão  $\mathbf{Ix1}$  que nos dá a contribuição relativa dos perfis linhas no K-ésimo eixo. Cada componente deste vetor nos mostra o quanto cada perfil linhas contribui da sua inércia individual, para a composição da inércia principal deste eixo. Isso mostra o quão bem representado está cada ponto, neste eixo;  $\mathbf{cr}_k = (cr_{1k}, \dots, cr_{nk})$ .

$cr_{ik} = (r_i F_{ik}^2) / (\sum_k r_i F_{ik}^2) = F_{ik}^2 / F_i^2 = \cos^2 \theta$  onde  $\theta$  é o ângulo formado entre o eixo  $k$  e o vetor que aponta para o  $i$ -ésimo perfil linha.

## 21 - CÁLCULO DAS COORDENADAS PADRONIZADAS DOS PERFIS LINHA SUPLEMENTARES

$$\text{FPS} = \text{FS} * \text{DA} \quad \text{dim}(\text{SxK})$$

### III - PROCEDIMENTO PARA INCLUIR PONTOS-COLUNAS SUPLEMENTARES

#### 1 - ENTRADA DOS DADOS

$$\text{NSC} \quad \text{dim}(\text{IxZ})$$

**I**: nº de linhas

**Z**: nº de colunas

**NSC** é a matriz de dados suplementares de dimensão  $\text{IxZ}$ , sendo que **I** é o número de linhas e **Z** é o número de colunas que está sendo considerada como ponto suplementar.  $\text{NSC} = [n_{iz}]$ .

#### 2 - DEFINIÇÃO DE VETORES AUXILIARES

$$\text{HZ} = (1, \dots, 1)^T \quad \text{dim}(\text{Zx1})$$

**HZ** é um vetor de dimensão  $\text{Zx1}$  com todos os elementos iguais a 1 (um); ( $1_z$ ).

#### 3 - CÁLCULO DOS DE ELEMENTOS MARGINAIS E TOTAL DA MATRIZ DE DADOS SUPLEMENTARES

$$\text{NTCS} = (\text{HI})^T * \text{NSC} * \text{HZ} \quad \text{escalar}$$

**NTCS** é o número de elementos da matriz de dados; ( $n_z$ ).

$$\text{CMCS} = \text{NCS} * \text{HZ} \quad \text{dim}(\text{Ix1})$$

**CMCS** é o vetor de dimensão  $\text{Ix1}$  que nos dá a coluna marginal da matriz de dados suplementares, cada elemento desse vetor é a soma das colunas em cada linha; ( $n_s$ ).

$$\text{LMCS} = (\text{HI})^T * \text{NCS} \quad \text{dim}(\text{Zx1})$$

**LMCS** é o vetor de dimensão  $\text{Zx1}$  que nos dá a linha marginal da matriz de dados, onde cada elemento desse vetor é a soma das linha em cada colunas; ( $n_{s_j}$ ).

#### 4 - CÁLCULO DA MATRIZ DE CORRESPONDÊNCIA

$$\text{PCS} = (1/\text{NTCS}) * \text{NCS} \quad \text{dim}(\text{IxZ})$$

**PCS** é uma matriz de dimensão  $\text{IxZ}$  com os elementos  $p_{ij} = n_{ij}/n$ . Quando **N** é uma tabela de contingência, **P** é uma matriz de frequências relativas.

**6 - CÁLCULO DO VETOR DE MASSAS DAS COLUNAS**

$$CS = PCS * HZ \quad \text{dim}(I \times I)$$

**RS** é um vetor de dimensão  $I \times I$  com os elementos  $rs_i = n_s / ns$ .  $rs = (rs_1, \dots, rs_I)$ .

**10 - DEFINIÇÃO DE MATRIZES DIAGONAIS**

$$DCS = \text{diag}(CS) \quad \text{dim}(I \times I)$$

**DCS** é uma matriz diagonal de dimensão  $I \times I$  com os elementos da diagonal iguais aos elementos do vetor marginal **CS**;  $D_{cs}$ .

$$DICS = \text{inv}(DCS) \quad \text{dim}(I \times I)$$

**DICS** é a matriz diagonal de dimensão  $I \times I$  com os elementos da diagonal iguais ao inverso dos elementos do vetor marginal **RS**;  $D_{rs}^{-1}$ .

**12 - CÁLCULO DA MATRIZ DOS PERFIS COLUNA**

$$CCS = DICS * PCS \quad \text{dim}(I \times Z)$$

**CC** é uma matriz dos perfis coluna de dimensão  $I \times Z$ .  $CS = D_{rs}^{-1} PCS$ .

**22 - CÁLCULO DOS COORDENADAS PRINCIPAIS DOS PERFIS COLUNA SUPLEMENTARES**

$$GS = CCS * FP \quad \text{dim}(Z \times K)$$

**GS** = [GS1 | ... | GSK] é uma matriz de dimensão  $Z \times K$  onde nas  $K$  colunas temos  $K$  vetores que nos dão as coordenadas principais dos  $Z$  perfis coluna suplementares em cada eixo.

$$GS1 \quad \text{dim}(Z \times 1)$$

**GS1** é o vetor de dimensão  $Z \times 1$  que nos dá as coordenadas (ou projeções) dos  $Z$  perfis coluna suplementares no eixo de maior inércia da nuvem.  $GS1 = (gs_{11}, \dots, gs_{I1})$ .

$$GSK \quad \text{dim}(Z \times 1)$$

**GSK** é o vetor de dimensão  $Z \times 1$  que nos dá as coordenadas (ou projeções) dos  $S$  perfis coluna suplementares no eixo de menor inércia ( $K$ -ésimo eixo) da nuvem.  $Gk = (gs_{1k}, \dots, gs_{Ik})$ .

**23 - CÁLCULO DE VETORES COM ELEMENTOS AO QUADRADO**

$$\mathbf{DGS1} = \text{diag}(\mathbf{GS1}) \quad \text{dim}(\mathbf{ZxZ})$$

$$\mathbf{DQGS1} = \mathbf{DGS1} * \mathbf{DGS1} \quad \text{dim}(\mathbf{ZxZ})$$

$$\mathbf{QGS1} = \mathbf{DQGS1} * \mathbf{HZ} \quad \text{dim}(\mathbf{Zx1})$$

$\mathbf{QGS1}$  é o vetor de dimensão  $\mathbf{Zx1}$  que nos dá o quadrado de cada componente do vetor de coordenadas  $\mathbf{GS1}$ .  $\mathbf{GS1}^2 = (gs_{11}^2, \dots, gs_{11}^2)$ .

$$\mathbf{DGSK} = \text{diag}(\mathbf{GSK}) \quad \text{dim}(\mathbf{ZxZ})$$

$$\mathbf{DQGSK} = \mathbf{DGSK} * \mathbf{DGSK} \quad \text{dim}(\mathbf{ZxZ})$$

$$\mathbf{QGSK} = \mathbf{DQGSK} * \mathbf{HZ} \quad \text{dim}(\mathbf{Zx1})$$

$\mathbf{QGSK}$  é o vetor de dimensão  $\mathbf{Zx1}$  que nos dá o quadrado de cada componente do vetor de coordenadas  $\mathbf{GSK}$ .  $\mathbf{GSK}^2 = (gs_{1k}^2, \dots, gs_{1k}^2)$

$$\mathbf{SQGS} = \mathbf{QGS1} + \dots + \mathbf{QGSK} \quad \text{dim}(\mathbf{Zx1})$$

$$\mathbf{DSQGS} = \text{diag}(\mathbf{SQGS}) \quad \text{dim}(\mathbf{ZxZ})$$

**26 - CÁLCULO DA CONTRIBUIÇÃO RELATIVA DOS PERFIS COLUNA SUPLEMENTARES EM CADA EIXO**

$$\mathbf{CRCS1} = \text{inv}(\mathbf{DSQGS}) * \mathbf{QGS1} \quad \text{dim}(\mathbf{Zx1})$$

$\mathbf{CRCS1}$  é o vetor de dimensão  $\mathbf{Zx1}$  que nos dá a contribuição relativa dos perfis coluna suplementares no primeiro eixo. Cada componente deste vetor nos mostra o quanto cada perfil coluna suolementar esta bem representado no primeiro eixo.  $\mathbf{cr}_1 = (cr_{11}, \dots, cr_{11})$ .

$$cr_{i1} = [In_c(\mathbf{R}_i)]_1 / \sum_k [In_c(\mathbf{R}_i)]_k = [In_c(\mathbf{R}_i)]_1 / In_c(\mathbf{R}_i) = (r_i F_{i1}^2) / (\sum_k r_i F_{ik}^2)$$

$$\mathbf{CRCSK} = \text{inv}(\mathbf{DSQGS}) * \mathbf{QGSK} \quad \text{dim}(\mathbf{Zx1})$$

$\mathbf{CRCK}$  é o vetor de dimensão  $\mathbf{Zx1}$  que nos dá a contribuição relativa dos perfis colunas no  $K$ -ésimo eixo. Cada componente deste vetor nos mostra o quanto cada perfil coluna contribui da sua inércia individual, para a composição da inércia principal deste eixo. Isso mostra o quão bem representado está cada ponto, neste eixo.  $\mathbf{cr}_k = (cr_{1k}, \dots, cr_{1k})$ .

$cr_{ik} = (r_i F_{ik}^2) / (\sum_k r_i F_{ik}^2) = F_{ik}^2 / F_i^2 = \cos^2 \theta$  onde  $\theta$  é o ângulo formado entro o eixo  $k$  e o vetor que aponta para o  $i$ -ésimo perfil linha.

**27 - CÁLCULO DAS COORDENADAS PADRONIZADAS DOS PERFIS COLUNA SUPLEMENTARES**

$$\mathbf{GPS} = \mathbf{GS} * \mathbf{DA} \quad \text{dim}(\mathbf{ZxK})$$

**B.2 PROGRAMA NO IML/SAS DO ALGORITMO DA AC**

```

/* 1 - ENTRADA DOS DADOS*/

/*TABELA DE CONTINGÊNCIA*/
data datal;
  input local $ i0 i1 i2 i3 i4 in r0 r1 r2 r3 r4 r5 r6 r7 r8 r9 r10 rn zu zr;
  cards;
L71 1376 2937 947 924 3731 0 1362 117 0 520 235
    217 943 1119 1075 1738 1945 644 8854 1061
L72 17412 17308 1887 2314 3588 0 2599 102 1510 1105 1193
    3107 2711 7398 7491 6314 7479 1500 31047 11462
L73 12804 17426 895 1101 1398 100 396 1802 3815 2704 3203
    3186 3002 3610 3404 3700 2202 2700 28324 5400
L74 11134 34332 6388 4131 3212 411 306 2760 3676 3906 4892
    4967 5002 7166 8912 7334 4126 6561 50127 9481
L75 139196 773127 300640 165657 84383 740 5779 141192 147193 156616 143004
    141287 143983 124585 125458 119540 88849 123874 1152049 311694
L81 5239 13379 7469 6490 3717 116 16 344 181 556 1284
    1389 3602 6050 7315 7956 7485 232 34593 1817
L82 17729 28753 7948 3899 5447 90 1033 663 5213 2113 5233
    5557 6042 8093 8176 8164 13073 506 55341 8525
L83 17032 21996 3851 1646 1597 28 348 4739 4613 4074 4165
    4987 4723 4404 4743 4556 4268 530 40577 5573
L84 17343 66950 36209 16647 4954 716 479 6611 7812 10535 12001
    12972 17379 22104 21933 18589 10268 2136 133178 9641
L85 113980 679286 787356 387996 95207 9731 3411 209603 218893 246437 238130
    223176 224880 206609 185872 165652 130123 20770 1833664 239892
L91 10344 42139 79224 40694 11066 271 1917 14187 15282 17956 17993
    19530 19049 18426 16976 18554 16804 7065 170980 12760
L92 17367 39459 17612 12457 15109 10 3411 2373 8705 3634 4710
    9659 6489 9840 11184 13906 26550 1554 93086 8929
L93 3330 8830 4685 2711 1706 21 436 1452 1535 1731 1874
    2096 1660 2248 2726 2599 2376 549 20204 1080
L94 5632 21918 45982 22336 6099 162 913 6887 8102 10040 9644
    11469 10200 10509 11233 10610 8319 4202 95749 6379
L95 132670 567576 1053442 482949 119372 3323 18593 195481 209893 247226 249694
    256136 238735 242994 229091 210788 165260 95442 2184566 174766
;
run;

/* I - Análise de Correspondência da Tabela de Contingência N */
proc iml;

use datal;
read all var {i0 i1 i2 i3 i4 in r0 r1 r2 r3 r4 r5 r6 r7 r8 r9 r10 rn zu zr}
into N;
title ' TABELA DE CONTINGÊNCIA ';
print N;

/* 2 - DEFINIÇÃO DE VETORES AUXILIARES*/

HI = J(15,1,1);
HJ = J(20,1,1);
HK = J(14,1,1);

/* 3 - CÁLCULO DOS ELEMENTOS MARGINAIS E TOTAL DA MATRIZ DE DADOS*/
title ' COLUNA MARGINAL ';

```

```
CM = N*HJ;

title ' LINHA MARGINAL ';
LM = t(N)*HI;

title ' NUMERO TOTAL ';
NT = t(HI)*N*HJ;

print NT,CM,LM;

/* 4 - CÁLCULO DA MATRIZ DE CORRESPONDÊNCIA*/
title ' MATRIZ DE CORRESPONDÊNCIA ';

P = (1/NT)*N;
print P;

/* 5 - CÁLCULO DO VETOR DE MASSAS DAS LINHAS*/
title ' VETOR DE MASSAS DAS LINHAS ';

R = P*HJ;
print R;

/* 6 - CÁLCULO DO VETOR DE MASSAS DAS COLUNAS*/
title ' VETOR DE MASSAS DAS COLUNAS ';

C = T(P)*HI;
print C;

/* 10 - DEFINIÇÃO DE MATRIZES DIAGONAIS*/

DR = diag(R);
DIR = inv(DR);
DRIR = sqrt(DIR);
DC = diag(C);
DIC = inv(DC);
DRIC = sqrt(DIC);

/* 11 - CÁLCULO DA MATRIZ DOS PERFIS LINHAS*/
title ' MATRIZ DE PERFIS DAS LINHA ';

RR = DIR*P;
print RR;

/* 12 - CÁLCULO DA MATRIZ DOS PERFIS COLUNAS*/
title ' MATRIZ DE PERFIS DAS COLUNA ';

CC = DIC*t(P);
TCC =t(CC);
print TCC;

/* 13 - CÁLCULO DA MATRIZ DOS PERFIS LINHAS CENTRALIZADO, PONDERADO E
TRANSFORMADO*/

X = DRIR*(P - R*t(C))*DRIC;
print X;

/* 14 - CALCULO DA MATRIZ DAS SOMAS DE QUADRADOS*/

S = t(X)*X;
print S;
Z = X*t(X);
```

```
print Z;

/* 14 - DECOMPOSIÇÃO EM VALORES SINGULARES DE X */
title ' DECOMPOSIÇÃO EM VALORES SINGULARES ';

call svd (U,A,V,X);
print U,A,V;

rank = 14;

UU =
U[,1]||U[,2]||U[,3]||U[,4]||U[,5]||U[,6]||U[,7]||U[,8]||U[,9]||U[,10]||U[,11]||
|U[,12]||U[,13]||U[,14];
print UU;

VV =
V[,1]||V[,2]||V[,3]||V[,4]||V[,5]||V[,6]||V[,7]||V[,8]||V[,9]||V[,10]||V[,11]||
|V[,12]||V[,13]||V[,14];
print VV;

AA = A[1:rank,];
print AA;

DA = diag(AA);

/* 15 - CÁLCULO DAS COORDENADAS PRINCIPAIS DOS PERFIS LINHAS*/
title ' COORDENADAS PRINCIPAIS DOS PERFIS LINHA ';

F = DRIR*UU*DA;
print F;

/* 21 - CÁLCULO DAS COORDENADAS PADRONIZADAS DOS PERFIS LINHAS*/
title ' COORDENADAS PADRONIZADAS DOS PERFIS LINHA ';

FP = F*inv(DA);
print FP;

/* 16 - CÁLCULO DE VETORES COM ELEMENTOS AO QUADRADO*/

DF1 = diag(F[,1]);
DQF1 = DF1*DF1;
QF1 = DQF1*HI;

DF2 = diag(F[,2]);
DQF2 = DF2*DF2;
QF2 = DQF2*HI;

DF3 = diag(F[,3]);
DQF3 = DF3*DF3;
QF3 = DQF3*HI;

DF4 = diag(F[,4]);
DQF4 = DF4*DF4;
QF4 = DQF4*HI;

DF5 = diag(F[,5]);
DQF5 = DF5*DF5;
QF5 = DQF5*HI;

DF6 = diag(F[,6]);
DQF6 = DF6*DF6;
```

```
QF6 = DQF6*HI;

DF7 = diag(F[,7]);
DQF7 = DF7*DF7;
QF7 = DQF7*HI;

DF8 = diag(F[,8]);
DQF8 = DF8*DF8;
QF8 = DQF8*HI;

DF9 = diag(F[,9]);
DQF9 = DF9*DF9;
QF9 = DQF9*HI;

DF10 = diag(F[,10]);
DQF10 = DF10*DF10;
QF10 = DQF10*HI;

DF11 = diag(F[,11]);
DQF11 = DF11*DF11;
QF11 = DQF11*HI;

DF12 = diag(F[,12]);
DQF12 = DF12*DF12;
QF12 = DQF12*HI;

DF13 = diag(F[,13]);
DQF13 = DF13*DF13;
QF13 = DQF13*HI;

DF14 = diag(F[,14]);
DQF14 = DF14*DF14;
QF14 = DQF14*HI;

DL = DA*DA;
L = DL*HK;
print L;

/* 17 - CÁLCULO DA COMPONENTE DE INÉRCIA DE CADA PERFIL LINHA EM CADA EIXO*/
title ' DECOMPOSIÇÃO DA INÉRCIA DOS PERFIS LINHA ';

CIR1 = DR*QF1;
CIR2 = DR*QF2;
CIR3 = DR*QF3;
CIR4 = DR*QF4;
CIR5 = DR*QF5;
CIR6 = DR*QF6;
CIR7 = DR*QF7;
CIR8 = DR*QF8;
CIR9 = DR*QF9;
CIR10 = DR*QF10;
CIR11 = DR*QF11;
CIR12 = DR*QF12;
CIR13 = DR*QF13;
CIR14 = DR*QF14;
CIR =
CIR1||CIR2||CIR3||CIR4||CIR5||CIR6||CIR7||CIR8||CIR9||CIR10||CIR11||CIR12||CIR
13||CIR14;
print CIR;

/* 19 - CÁLCULO DA CONTRIBUIÇÃO RELATIVA DOS PERFIS LINHA EM CADA EIXO*/
```

```

title ' CONTRIBUIÇÃO RELATIVA DOS PERFIS LINHA ';

CRR1 = inv(DIR)*CIR1;
CRR2 = inv(DIR)*CIR2;
CRR3 = inv(DIR)*CIR3;
CRR4 = inv(DIR)*CIR4;
CRR5 = inv(DIR)*CIR5;
CRR6 = inv(DIR)*CIR6;
CRR7 = inv(DIR)*CIR7;
CRR8 = inv(DIR)*CIR8;
CRR9 = inv(DIR)*CIR9;
CRR10 = inv(DIR)*CIR10;
CRR11 = inv(DIR)*CIR11;
CRR12 = inv(DIR)*CIR12;
CRR13 = inv(DIR)*CIR13;
CRR14 = inv(DIR)*CIR14;
CRR =
CRR1||CRR2||CRR3||CRR4||CRR5||CRR6||CRR7||CRR8||CRR9||CRR10||CRR11||CRR12||CRR
13||CRR14;
print CRR;

/* 20 - CÁLCULO DA CONTRIBUIÇÃO ABSOLUTA DE CADA PERFIL LINHA EM CADA EIXO*/
title ' CONTRIBUIÇÃO ABSOLUTA DOS PERFIS LINHA ';

CAR1 =(1/L[1])*CIR1;
CAR2 =(1/L[2])*CIR2;
CAR3 =(1/L[3])*CIR3;
CAR4 =(1/L[4])*CIR4;
CAR5 =(1/L[5])*CIR5;
CAR6 =(1/L[6])*CIR6;
CAR7 =(1/L[7])*CIR7;
CAR8 =(1/L[8])*CIR8;
CAR9 =(1/L[9])*CIR9;
CAR10 =(1/L[10])*CIR10;
CAR11 =(1/L[11])*CIR11;
CAR12 =(1/L[12])*CIR12;
CAR13 =(1/L[13])*CIR13;
CAR14 =(1/L[14])*CIR14;
CAR =
CAR1||CAR2||CAR3||CAR4||CAR5||CAR6||CAR7||CAR8||CAR9||CAR10||CAR11||CAR12||CAR
13||CAR14;
print CAR;

/* 22 - CÁLCULO DAS COORDENADAS PRINCIPAIS DOS PERFIS COLUNAS*/
title ' COORDENADAS PRINCIPAIS DOS PERFIS COLUNA ';

G = DRIC*VV*DA;
print G;

/* 28 - CÁLCULO DAS COORDENADAS PADRONIZADAS DOS PERFIS COLUNAS*/
title ' COORDENADAS PADRONIZADAS DOS PERFIS COLUNA';

GP = G*inv(DA);
print GP;
/* 22 - CÁLCULO DE VETORES COM ELEMENTOS AO QUADRADO*/

DG1 = diag(G[,1]);
DQG1 = DG1*DG1;
QG1 = DQG1*HJ;

DG2 = diag(G[,2]);

```

```
DQG2 = DG2*DG2;  
QG2 = DQG2*HJ;
```

```
DG3 = diag(G[,3]);  
DQG3 = DG3*DG3;  
QG3 = DQG3*HJ;
```

```
DG4 = diag(G[,4]);  
DQG4 = DG4*DG4;  
QG4 = DQG4*HJ;
```

```
DG5 = diag(G[,5]);  
DQG5 = DG5*DG5;  
QG5 = DQG5*HJ;
```

```
DG6 = diag(G[,6]);  
DQG6 = DG6*DG6;  
QG6 = DQG6*HJ;
```

```
DG7 = diag(G[,7]);  
DQG7 = DG7*DG7;  
QG7 = DQG7*HJ;
```

```
DG8 = diag(G[,8]);  
DQG8 = DG8*DG8;  
QG8 = DQG8*HJ;
```

```
DG9 = diag(G[,9]);  
DQG9 = DG9*DG9;  
QG9 = DQG9*HJ;
```

```
DG10 = diag(G[,10]);  
DQG10 = DG10*DG10;  
QG10 = DQG10*HJ;
```

```
DG11 = diag(G[,11]);  
DQG11 = DG11*DG11;  
QG11 = DQG11*HJ;
```

```
DG12 = diag(G[,12]);  
DQG12 = DG12*DG12;  
QG12 = DQG12*HJ;
```

```
DG13 = diag(G[,13]);  
DQG13 = DG13*DG13;  
QG13 = DQG13*HJ;
```

```
DG14 = diag(G[,14]);  
DQG14 = DG14*DG14;  
QG14 = DQG14*HJ;
```

```
/* 24 - CÁLCULO DA COMPONENTE DE INÉRCIA DE CADA PERFIL COLUNA EM CADA EIXO*/  
title ' DECOMPOSIÇÃO DA INÉRCIA DOS PERFIS COLUNA';
```

```
CIC1 = DC*QG1;  
CIC2 = DC*QG2;  
CIC3 = DC*QG3;  
CIC4 = DC*QG4;  
CIC5 = DC*QG5;  
CIC6 = DC*QG6;  
CIC7 = DC*QG7;
```

```
CIC8 = DC*QG8;
CIC9 = DC*QG9;
CIC10 = DC*QG10;
CIC11 = DC*QG11;
CIC12 = DC*QG12;
CIC13 = DC*QG13;
CIC14 = DC*QG14;
CIC =
CIC1||CIC2||CIC3||CIC4||CIC5||CIC6||CIC7||CIC8||CIC9||CIC10||CIC11||CIC12||CIC
13||CIC14;
print CIC;

/* 26 - CÁLCULO DA CONTRIBUIÇÃO RELATIVA DOS PERFIS COLUNA EM CADA EIXO*/
title ' CONTRIBUIÇÃO RELATIVA DOS PERFIS COLUNA';

CRC1 = inv(DIC)*CIC1;
CRC2 = inv(DIC)*CIC2;
CRC3 = inv(DIC)*CIC3;
CRC4 = inv(DIC)*CIC4;
CRC5 = inv(DIC)*CIC5;
CRC6 = inv(DIC)*CIC6;
CRC7 = inv(DIC)*CIC7;
CRC8 = inv(DIC)*CIC8;
CRC9 = inv(DIC)*CIC9;
CRC10 = inv(DIC)*CIC10;
CRC11 = inv(DIC)*CIC11;
CRC12 = inv(DIC)*CIC12;
CRC13 = inv(DIC)*CIC13;
CRC14 = inv(DIC)*CIC14;
CRC =
CRC1||CRC2||CRC3||CRC4||CRC5||CRC6||CRC7||CRC8||CRC9||CRC10||CRC11||CRC12||CRC
13||CRC14;
print CRC;

/* 27 - CÁLCULO DA CONTRIBUIÇÃO ABSOLUTA DOS PERFIS COLUNA EM CADA EIXO*/
title ' CONTRIBUIÇÃO ABSOLUTA DOS PERFIS COLUNA';

CAC1 = (1/L[1])*CIC1;
CAC2 = (1/L[2])*CIC2;
CAC3 = (1/L[3])*CIC3;
CAC4 = (1/L[4])*CIC4;
CAC5 = (1/L[5])*CIC5;
CAC6 = (1/L[6])*CIC6;
CAC7 = (1/L[7])*CIC7;
CAC8 = (1/L[8])*CIC8;
CAC9 = (1/L[9])*CIC9;
CAC10 = (1/L[10])*CIC10;
CAC11 = (1/L[11])*CIC11;
CAC12 = (1/L[12])*CIC12;
CAC13 = (1/L[13])*CIC13;
CAC14 = (1/L[14])*CIC14;
CAC =
CAC1||CAC2||CAC3||CAC4||CAC5||CAC6||CAC7||CAC8||CAC9||CAC10||CAC11||CAC12||CAC
13||CAC14;
print CAC;
```



CENTRÓIDE DAS LINHAS OU VETOR DE MASSAS DAS COLUNAS

C
0.0258451
0.1155002
0.1164458
0.0569709
0.0178331
0.0007774
0.0020276
0.0290956
0.0314749
0.0350718
0.0344834
0.0346061
0.0340455
0.0333904
0.0319282
0.0296736
0.0241902
0.0132673
0.2933895
0.0399832

MATRIZ DOS PERFIS LINHA

RR

0.045599 0.087939 0.081873 0.081094 0.125438 0.045892 0.038834 0.0704919 0.079005 0.007295 0.037028 0.0376198 0.0361405 0.05943 0.055821 0.021607 0.296655 0.038877
0.135578 0.135208 0.047989 0.0281482 0.0281482 0.030289 0.007999 0.018406 0.008668 0.008599 0.028635 0.021252 0.0380112 0.0384405 0.0495111 0.0386624 0.017622 0.243643 0.038877
0.128558 0.172243 0.008965 0.008925 0.0138181 0.000994 0.009341 0.018113 0.037081 0.026228 0.038659 0.0314919 0.029522 0.0386818 0.0366457 0.0365714 0.021649 0.0088972 0.279989 0.055384
0.026223 0.191977 0.035223 0.0281019 0.0179618 0.002298 0.017112 0.015642 0.002555 0.021427 0.027365 0.027759 0.027976 0.040729 0.0499367 0.040124 0.023073 0.0086897 0.2831847 0.0550184
0.017159 0.176572 0.048509 0.037745 0.0192257 0.001486 0.013167 0.032706 0.03558 0.03865 0.028285 0.029193 0.038046 0.038867 0.028876 0.027372 0.020243 0.0282847 0.2524947 0.0710194
0.04793 0.122847 0.048387 0.0354159 0.043021 0.01032 0.001465 0.038438 0.0016571 0.030302 0.011755 0.0127163 0.032943 0.035877 0.036988 0.072871 0.048251 0.002124 0.318887 0.016694
0.025223 0.151094 0.0414827 0.023299 0.028223 0.0004897 0.005815 0.008434 0.02228 0.01028 0.0273124 0.029104 0.031548 0.042295 0.042527 0.04281 0.048234 0.028609 0.288881 0.0444942
0.123018 0.158732 0.0276151 0.018889 0.012549 0.000222 0.002515 0.034229 0.033149 0.029258 0.000381 0.036012 0.034134 0.038403 0.0342579 0.039072 0.03827 0.038281 0.293805 0.042507
0.040478 0.152584 0.048102 0.038534 0.011524 0.011671 0.01118 0.015249 0.018239 0.024582 0.02009 0.030281 0.040618 0.038588 0.0511907 0.048359 0.023451 0.004858 0.310817 0.022507
0.038228 0.129482 0.123571 0.023271 0.015049 0.001563 0.001563 0.038394 0.038394 0.03839 0.038649 0.038805 0.038785 0.038505 0.038213 0.029797 0.028223 0.020278 0.038889 0.294786 0.038657
0.018457 0.018472 0.143726 0.078257 0.070745 0.000486 0.003478 0.025376 0.027241 0.028552 0.036623 0.036607 0.048881 0.038279 0.037973 0.038801 0.030485 0.018271 0.310884 0.028441
0.056467 0.128824 0.055473 0.040783 0.048887 0.000827 0.011465 0.007538 0.028436 0.018741 0.015889 0.038608 0.021208 0.032152 0.036408 0.046379 0.036522 0.030777 0.301889 0.029075
0.025243 0.138835 0.073782 0.024836 0.028718 0.000389 0.000386 0.022412 0.020471 0.021108 0.022905 0.038275 0.025889 0.038281 0.042846 0.040784 0.032228 0.038594 0.316041 0.028441
0.038821 0.071574 0.151079 0.072917 0.019063 0.001567 0.002979 0.022478 0.028439 0.028282 0.0384767 0.038433 0.038284 0.0383 0.03863 0.038286 0.0271521 0.037143 0.312512 0.038802
0.018744 0.03868 0.148833 0.048234 0.018652 0.001465 0.001465 0.027681 0.027681 0.02768 0.038275 0.038276 0.038272 0.038309 0.038286 0.028607 0.028384 0.038483 0.308818 0.028441

MATRIZ TRANSPOSTA DOS PERFIS COLUNA

TCC

0.00288 0.001256 0.00022 0.001821 0.01097 0.008228 0.001989 0.000733 0.00037 0.00301 0.001388 0.001654 0.001661 0.002867 0.008945 0.002006 0.001495 0.00132
0.038388 0.007411 0.008104 0.00208 0.008815 0.005398 0.001734 0.002326 0.001582 0.00171 0.004478 0.008981 0.00195 0.011684 0.00228 0.015295 0.008815 0.00235 0.04177
0.02801 0.007466 0.00181 0.00185 0.00187 0.00637 0.00868 0.0016 0.03944 0.0083 0.004587 0.004552 0.00348 0.00349 0.00227 0.001667 0.004019 0.001667 0.00745 0.006594
0.02105 0.04706 0.02713 0.00881 0.00977 0.02847 0.00466 0.00434 0.00576 0.0088 0.00701 0.00704 0.00728 0.00659 0.00304 0.02223 0.00454 0.02452 0.00848 0.01222
0.26659 0.33047 0.12788 0.14885 0.23016 0.07078 0.140547 0.299947 0.23027 0.22894 0.23027 0.21915 0.21915 0.3848 0.194311 0.19223 0.181681 0.43738 0.191081 0.38540
0.00251 0.005287 0.00172 0.00639 0.000382 0.007346 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803
0.038534 0.02317 0.038535 0.008847 0.01506 0.05256 0.021957 0.00127 0.008191 0.022955 0.00783 0.007916 0.00789 0.001989 0.002884 0.013867 0.03272 0.008882 0.03887 0.00547
0.02926 0.0048 0.01686 0.001489 0.00449 0.00173 0.00848 0.00832 0.00728 0.005449 0.005934 0.00727 0.00888 0.006229 0.007348 0.00738 0.00828 0.001957 0.0084 0.00894
0.03888 0.028623 0.053784 0.044511 0.03787 0.0465 0.01482 0.01232 0.02249 0.04488 0.012718 0.008584 0.02545 0.02721 0.00836 0.00887 0.00288 0.007923 0.02848 0.01923
0.2808 0.2983 0.39498 0.38814 0.26081 0.81959 0.08192 0.3628 0.31926 0.37519 0.3155 0.38926 0.38505 0.30071 0.27807 0.27807 0.27807 0.07424 0.30963 0.296271
0.01978 0.0045 0.03874 0.03821 0.00889 0.00728 0.04552 0.02047 0.02012 0.02538 0.02815 0.02916 0.0274 0.02725 0.02854 0.00233 0.03851 0.02859 0.00827 0.01581
0.03827 0.00889 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803 0.001803
0.00521 0.00889 0.001989 0.00254 0.00732 0.00136 0.00164 0.00248 0.002419 0.00249 0.00287 0.00294 0.0024 0.00396 0.00225 0.00437 0.00456 0.00245 0.00457 0.00135
0.00771 0.00881 0.01924 0.01887 0.01841 0.01036 0.02288 0.011784 0.02735 0.04157 0.01384 0.01636 0.01487 0.01563 0.01739 0.01783 0.00709 0.01666 0.01642 0.00802
0.25871 0.26029 0.44708 0.41924 0.3305 0.21402 0.45489 0.32278 0.3801 0.36525 0.381 0.36047 0.34599 0.38985 0.34888 0.38133 0.38683 0.35751 0.38247 0.21671

MATRIZ DOS PERFIS LINHAS CENTRALIZADO, PONDERADO E TRANSFORMADO

X

0.02875 -0.00882 -0.0081 -0.00163 0.00089 -0.00089 0.02726 -0.0088 -0.0088 -0.0082 -0.0082 -0.0081 -0.0087 0.00887 0.00092 0.00607 0.00197 0.00295 0.00026 -0.0082
0.05885 0.00725 -0.02857 -0.01298 0.00266 -0.0224 0.02263 -0.01374 -0.0889 -0.01189 -0.01047 -0.0243 -0.00804 0.00704 0.01967 0.00846 0.02837 -0.00028 -0.0281 0.01889
0.048129 0.018899 -0.02204 -0.01359 -0.02217 0.001553 0.002934 -0.0048 0.00282 -0.0082 -0.0076 -0.00185 -0.00167 0.00887 0.00159 0.00835 -0.00103 0.008213 -0.00154 0.004732
0.02109 0.02161 -0.0226 -0.01345 0.00096 0.00513 -0.0086 -0.00582 -0.0058 0.00863 -0.00869 -0.00583 -0.00806 0.00481 0.00423 0.00892 -0.00166 0.01228 -0.0227 0.00802
0.00704 0.00852 -0.0058 -0.0082 -0.00073 -0.00073 0.00889 0.00477 0.00854 -0.00475 -0.00845 -0.00828 -0.01257 -0.00825 -0.00828 -0.01257 -0.00825 -0.00828 -0.01257 0.02812
0.00112 0.001505 -0.00085 0.00029 0.00884 0.00082 -0.00071 -0.0018 -0.0225 -0.0127 -0.00885 -0.00889 -0.0026 0.00889 0.00413 0.00816 0.00881 -0.00111 0.00889 -0.00882
0.0082 -0.00805 -0.00804 -0.0085 0.00724 -0.00074 0.002719 -0.00489 -0.00281 -0.01286 -0.00889 -0.0082 -0.0025 0.00714 0.00883 0.007303 0.02816 -0.0088 -0.0088 0.00281
0.00109 0.00885 -0.0242 -0.00829 -0.00805 -0.0007 0.00889 0.00281 0.00881 -0.0028 -0.0088 0.00889 0.00885 -0.00716 0.0088 0.001553 0.00881 -0.00881 -0.00047 0.00116
0.01325 0.00757 -0.0132 -0.0009 -0.00885 0.00489 -0.00281 -0.0182 -0.00885 -0.00849 -0.00075 -0.0088 0.00549 0.00488 0.005823 0.01585 -0.0021 -0.00467 0.00485 -0.01226
-0.0281 -0.00885 0.01657 0.01581 -0.0001 0.00541 -0.00822 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488 0.00488
-0.0021 -0.00883 0.01398 0.00882 0.02726 -0.00182 0.00571 -0.0082 -0.00881 -0.0021 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881
0.0088 0.00885 -0.0225 -0.00885 0.00889 -0.00885 0.008813 -0.0158 -0.00202 -0.01589 -0.0125 -0.0014 -0.00853 -0.00884 0.00879 0.01258 0.00881 -0.00887 0.00881 -0.0088
0.00881 0.00889 -0.0082 -0.00816 0.00883 -0.00884 0.008813 -0.0088 -0.0225 -0.00889 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881 -0.00881
-0.00884 -0.00883 0.01235 0.00889 0.00881 -0.00889 0.00882 -0.0075 -0.00881 -0.00884 -0.00889 0.00883 -0.00883 0.008827 0.00888 0.008816 0.00884 0.00882 0.00888 -0.01179
-0.00884 -0.00884 0.00884 0.00884 -0.00884 0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884 -0.00884



VALORES SINGULARES

AA

0.2077744  
 0.1319346  
 0.0731337  
 0.0573244  
 0.0478844  
 0.0325114  
 0.0248714  
 0.0136564  
 0.0107247  
 0.0081797  
 0.0052999  
 0.0033584  
 0.0030905  
 0.0024687

MATRIZ DAS COORDENADAS PRINCIPAIS DOS PERFIS LINHA

F

0.3544965 0.8242045 0.7164671 -0.432703 0.2667015 0.4622801 0.3089921 0.0745622 -0.001175 -0.06303 -0.00954 0.0059368 0.0024552 -0.000944  
 0.6783117 0.6593659 -0.02285 0.1149894 0.2229701 0.177333 -0.186303 -0.016348 -0.004408 0.0236556 0.0008414 -0.000995 -0.001254 -0.00014  
 0.6104882 0.255334 -0.186721 0.2896664 0.1240713 -0.049997 0.090714 0.0413644 0.0515282 -0.015665 0.0351997 -0.022546 -0.002793 -0.006108  
 0.4546182 0.0723604 0.0056465 0.1414158 -0.143963 0.0331955 0.0321612 0.0140714 0.0782367 0.0226552 -0.017876 0.0131503 -0.001865 0.0008372  
 0.2965126 -0.130912 0.0452888 -0.003991 -0.004142 -0.004983 -0.003228 -0.000314 -0.008847 -0.000772 0.0003285 -0.00022 0.0001645 0.0000692  
 0.1250581 0.4686141 -0.040316 -0.092489 -0.334709 -0.006818 -0.09578 0.1262176 -0.026317 -0.010836 -0.008136 -0.008234 -0.004014 -0.002751  
 0.4069893 0.4584361 -0.135817 -0.030211 -0.000888 -0.090382 -0.032103 -0.035107 0.0144399 -0.057854 -0.008496 0.003842 -0.000012 0.0089611  
 0.4725719 0.2989886 -0.321492 0.1946771 0.150635 -0.081034 0.1018439 0.0374653 -0.069747 0.0139222 -0.014934 0.0140892 0.0061579 -0.002725  
 0.1167536 0.136215 -0.142758 0.0408925 -0.188258 0.0875304 0.0350666 -0.041743 -0.016913 -0.000901 0.0068301 0.0003119 0.0012319 -0.002484  
 -0.084193 -0.045318 -0.073103 -0.045061 0.016191 0.0090509 -0.001127 0.0026498 0.003604 0.0006507 -0.000075 -0.000032 0.0000133 0.0000244  
 -0.164771 0.081095 0.066251 0.0055686 -0.006068 -0.01711 -0.006983 0.0144196 -0.003368 0.0026648 0.0211817 0.011052 -0.004627 0.0043533  
 0.2440235 0.5078197 0.1037201 -0.232902 -0.006935 -0.111647 0.0016544 -0.024623 0.0067717 0.00235447 0.0028931 -0.001516 0.0007945 -0.005673  
 0.14797 0.2532717 -0.026924 -0.00539 -0.040446 0.0133861 0.1116554 -0.007114 -0.023438 0.0519199 -0.011286 -0.023411 -0.012269 0.0327585  
 -0.185236 0.0853183 0.0689239 0.0336212 -0.027213 -0.003938 -0.011585 0.0117802 0.006804 0.0049417 0.0059836 -0.001307 0.0228691 0.0051205  
 -0.16408 0.0321954 0.0429556 0.0383028 0.0010179 -0.005079 0.000275 -0.002807 -0.001317 -0.000996 -0.001983 -0.000651 -0.000681 -0.000599

INÉRCIAS PRINCIPAIS

L

0.0431702  
 0.0174067  
 0.0053485  
 0.0032861  
 0.0022929  
 0.001057  
 0.0006186  
 0.0001865  
 0.000115  
 0.0000669  
 0.0000281  
 0.0000113  
 9.5514E-6  
 6.0943E-6

DECOMPOSIÇÃO DA INERCIA DOS PERFIS LINHA

CMR

0.0001849 0.0009993 0.0007551 0.0002754 0.0001046 0.0003144 0.0001405 8.1784E-6 2.0303E-9 5.8443E-6 1.3389E-7 5.1849E-8 8.9674E-9 1.3107E-9  
 0.0029019 0.002742 3.5876E-6 0.0000834 0.0003136 0.0001983 0.0002189 1.6856E-6 1.2253E-7 3.5293E-6 4.4655E-9 6.239E-9 9.9245E-9 1.231E-10  
 0.0018648 0.0003262 0.0001744 0.0004198 0.0000777 0.0000125 0.0000412 8.5611E-6 0.0000133 1.2278E-6 6.1995E-6 2.5434E-6 3.9025E-8 1.8669E-7  
 0.0018278 0.0000463 2.8197E-7 0.0001769 0.0001833 9.7443E-6 9.1477E-6 1.7511E-6 0.0000541 4.5392E-6 2.8261E-6 1.5294E-6 3.0801E-8 6.1986E-9  
 0.0190834 0.0037198 0.0004452 3.4577E-6 3.7233E-6 5.3892E-6 2.2614E-6 2.1399E-8 3.2123E-6 1.2925E-7 2.3421E-8 1.0518E-8 5.8717E-9 1.0383E-9  
 0.0000845 0.0011863 8.7805E-6 0.0000462 0.0006052 2.5109E-7 0.0000496 0.0000861 3.7413E-6 6.3434E-7 3.5763E-7 3.663E-7 8.7042E-8 4.0884E-8  
 0.0015696 0.0019914 0.0001748 8.6483E-6 7.474E-9 0.0000774 9.7657E-6 0.0000117 1.9758E-6 0.0000317 6.8399E-7 1.3987E-7 1.278E-12 7.609E-7  
 0.0015291 0.0006121 0.0007077 0.0002595 0.0001554 0.000045 0.000071 9.611E-6 0.0000244 1.3272E-6 1.5271E-6 1.3592E-6 1.8216E-7 5.0851E-8  
 0.0002888 0.0003932 0.0004318 0.0000354 0.000751 0.0001623 0.0000261 0.0000369 6.0613E-6 1.7188E-8 1.6522E-6 2.0615E-9 3.2157E-8 1.3079E-7  
 0.0021808 0.0006318 0.0016441 0.0006247 0.0000807 0.0000252 3.9078E-7 2.1601E-6 3.9959E-6 1.3027E-7 1.7146E-9 3.212E-10 5.419E-11 1.827E-10  
 0.0007401 0.0001793 0.0001197 8.4534E-7 1.0039E-6 7.9803E-6 1.3294E-6 5.6682E-6 3.0917E-7 1.9359E-7 0.0000122 3.3299E-6 5.8374E-7 5.1663E-7  
 0.0009013 0.00039032 0.0001628 0.000821 7.2787E-7 0.0001887 7.0974E-6 9.1765E-6 6.9406E-7 8.3905E-6 1.2668E-6 3.4782E-8 9.5636E-9 4.8717E-7  
 0.0000691 0.0002026 2.289E-6 9.1745E-8 5.1655E-6 5.6574E-7 0.0000394 1.5979E-7 1.7347E-6 8.5122E-6 4.0219E-7 1.7306E-6 4.753E-7 3.3886E-6  
 0.0005199 0.0001103 0.000072 0.0000171 0.0000112 2.3501E-7 2.0336E-6 2.1028E-6 7.0148E-7 3.7003E-7 5.4251E-7 2.5877E-8 7.9247E-6 3.9729E-7  
 0.0094242 0.0003628 0.0006459 0.0005136 3.6268E-7 9.0287E-6 2.6471E-8 2.7585E-6 6.0689E-7 3.4695E-7 1.377E-6 1.4831E-7 1.6221E-7 1.2558E-7

INERCIA DOS PERFIS LINHA

IR  
 0.0027884  
 0.0064671  
 0.002948  
 0.0023183  
 0.0232666  
 0.0020721  
 0.0038786  
 0.0034183  
 0.0021335  
 0.0051939  
 0.001073  
 0.0060038  
 0.0003356  
 0.0007449  
 0.0109614

CONTRIBUIÇÃO RELATIVA DOS PERFIS LINHAS

CR  
 0.0662975 0.3583795 0.2708106 0.0987763 0.0375253 0.1127414 0.0503695 0.002933 7.2911E-7 0.0020959 0.000048 0.0000186 3.1801E-6 4.7006E-7  
 0.4487172 0.4240011 0.0005547 0.0128952 0.049485 0.0306686 0.0338497 0.0002606 0.0000189 0.0005457 6.905E-7 9.6474E-7 1.5346E-6 1.9042E-8  
 0.6325573 0.1106529 0.059174 0.1424104 0.0261269 0.0042425 0.0139667 0.0002904 0.0045065 0.0004165 0.0021029 0.0008628 0.0000132 0.0000633  
 0.7884419 0.0199746 0.0001216 0.0762908 0.0790634 0.0042032 0.0039459 0.0007554 0.0233505 0.001958 0.001219 0.0006597 0.0000133 2.6738E-6  
 0.820203 0.1598788 0.0191345 0.0001486 0.000016 0.0002316 0.0000972 9.1974E-7 0.0001381 5.5552E-6 1.0066E-6 4.5208E-7 2.5237E-7 4.4624E-8  
 0.0407739 0.5725183 0.0042376 0.0223019 0.2920737 0.0001212 0.023917 0.0415335 0.0018056 0.0003061 0.0001726 0.0001768 0.0000042 0.0000197  
 0.4046733 0.5134476 0.045066 0.0022299 1.9E7E-6 0.0199572 0.0005179 0.0030111 0.0006094 0.0061771 0.0001763 0.0000361 3.295E-10 0.0001962  
 0.4473397 0.1790651 0.2070346 0.0759157 0.0454521 0.0131534 0.0007765 0.0028116 0.0071505 0.0003883 0.0004467 0.0003976 0.0000533 0.0000149  
 0.1353663 0.1842824 0.2024099 0.0166081 0.3519957 0.0760943 0.012213 0.0173085 0.002841 8.0562E-6 0.0007744 9.662E-7 0.0000151 0.0000613  
 0.4196866 0.1216494 0.3165449 0.1202711 0.0155278 0.0048523 0.0000752 0.0004159 0.0007693 0.0000251 3.3012E-7 6.1839E-8 1.0433E-8 3.5175E-8  
 0.689738 0.1670758 0.1115089 0.0007878 0.0009365 0.0074371 0.0012389 0.0052824 0.0002881 0.0001804 0.0113984 0.0031032 0.000544 0.0004815  
 0.1501216 0.6501279 0.027121 0.1367497 0.0001212 0.0314251 0.0011822 0.0015285 0.0001156 0.0013975 0.0000211 5.7934E-6 1.5913E-6 0.0000811  
 0.2060286 0.6026057 0.006821 0.0002734 0.015393 0.0018859 0.1173113 0.0004782 0.0051692 0.0253658 0.0011985 0.0005152 0.0014164 0.0000979  
 0.6979875 0.1480756 0.0966357 0.0229945 0.0150643 0.0003155 0.0027301 0.0002823 0.0009417 0.0004968 0.0007283 0.0000347 0.0106389 0.0005334  
 0.859758 0.0331017 0.0589255 0.0468515 0.0000331 0.0008237 2.4149E-6 0.0002517 0.0000554 0.0000317 0.0001256 0.0000135 0.0000148 0.0000115

CONTRIBUIÇÃO ABSOLUTA DOS PERFIS LINHA

CR  
 0.0042823 0.0574097 0.1411856 0.0638171 0.0456346 0.2974203 0.2270524 0.0436529 0.0000177 0.0873491 0.0047665 0.0045971 0.0009284 0.0002151  
 0.0672196 0.1575276 0.0006708 0.025378 0.1367493 0.187641 0.3538833 0.0090381 0.0010653 0.0527488 0.000159 0.0005632 0.0010391 0.0000202  
 0.0431966 0.0187404 0.0326159 0.1277604 0.0335918 0.0118328 0.0665621 0.0459051 0.115505 0.0183507 0.2207068 0.2255095 0.0040857 0.0306336  
 0.0423403 0.0026603 0.0000527 0.0538222 0.0799383 0.0092189 0.014788 0.0093896 0.4706498 0.0678431 0.1006109 0.1356301 0.0032247 0.0010171  
 0.4420499 0.2137014 0.083237 0.0010522 0.0016238 0.0050986 0.0036557 0.0001147 0.0279288 0.0019318 0.0008338 0.0009326 0.0006147 0.0001704  
 0.001957 0.0681513 0.0016417 0.0140626 0.2639402 0.0002375 0.0801139 0.461455 0.0325276 0.0094808 0.0127319 0.0324768 0.0091113 0.0067086  
 0.0363573 0.1144063 0.0326803 0.0026318 3.2594E-6 0.0732317 0.0157871 0.0626226 0.0171779 0.4740186 0.0243503 0.012401 1.3382E-7 0.1248561  
 0.0354212 0.0351644 0.1323177 0.07897 0.0677602 0.0425379 0.1148103 0.0515345 0.2125086 0.0198361 0.0543649 0.1205115 0.0190714 0.0083441  
 0.0066909 0.022587 0.08074 0.0107829 0.3275248 0.153593 0.0421223 0.1979841 0.0526987 0.0002569 0.0588183 0.0001828 0.0033667 0.0214619  
 0.0505156 0.0562985 0.3073942 0.1900979 0.0351736 0.0238433 0.006317 0.0115826 0.0347414 0.001947 0.000061 0.0000285 5.6734E-6 0.00003  
 0.0171442 0.0102994 0.0223713 0.0002572 0.0004378 0.00755 0.0021491 0.0303893 0.002688 0.0028934 0.4354318 0.2982353 0.0611153 0.0847732  
 0.0208777 0.2242355 0.0304434 0.2498447 0.0003174 0.1784955 0.0114735 0.0492046 0.0060343 0.1254048 0.00451 0.0030839 0.0010002 0.0799887  
 0.0016015 0.0116367 0.000428 0.0000279 0.0022528 0.0005352 0.0636402 0.0008568 0.0150818 0.127223 0.0143183 0.1534439 0.0497622 0.556604  
 0.0120435 0.0063366 0.0134583 0.0052123 0.0048938 0.0002223 0.0032875 0.0112751 0.0060988 0.0055305 0.0193139 0.0022943 0.8298893 0.0651912  
 0.2183026 0.0208449 0.1207632 0.1562827 0.0001582 0.0085419 0.0000428 0.0147914 0.0052765 0.0051855 0.0490225 0.0131495 0.0169834 0.0206059

MATRIZ DAS COORDENADAS PADRONIZADAS DAS PERFIS LINHA

CP  
 2.4384912 2.9234217 -2.100903 3.0683627 2.2526783 -0.862305 0.1293159 1.1010679 0.0554041 -0.350543 -0.118086 0.158453 0.1896305 -0.09347  
 1.4825296 -0.744035 -0.364124 -0.61052 -0.94138 0.1627666 0.6496397 -0.817669 -0.317169 0.3526423 0.1755698 -0.043155 -0.281423 0.1985852  
 -1.48844 -0.114727 0.3288818 0.4911513 0.3061337 -0.053951 -0.570173 -0.649251 -0.141602 -0.48405 -0.820902 0.1490907 0.186174 -0.567574  
 -1.219065 -0.052211 0.3353602 -0.140577 0.1057463 -0.04806 -1.007495 1.5268549 0.2651467 1.0583952 1.9520965 -0.195503 -0.522809 1.1521164  
 0.5453632 1.5017472 2.5398641 -3.083402 0.5705496 0.3723564 2.4338663 2.8469111 0.7871445 -2.284707 -1.797914 -1.031949 2.1321564 -0.78209  
 -1.205962 -0.46742 -7.731199 -3.896642 -1.812913 7.3139876 2.3241851 4.8462294 27.315287 6.0161588 0.462016 15.764663 -1.976028 -7.266311  
 1.1621522 7.5292522 8.4978509 -2.210907 9.3210789 11.477561 3.1997186 -7.208154 -1.057615 4.614385 3.1006263 1.3347359 -4.024987 2.243974  
 -0.192388 -1.44249 -0.546737 -0.391293 1.4032709 -0.174499 0.8959485 2.0179946 -2.220791 2.1137266 -1.969979 1.4945622 -0.174508 -1.184982  
 -0.022957 -0.770708 -0.566487 -0.596691 1.4748949 -1.525676 0.026831 -0.787468 1.5462454 -0.0427 1.1145394 -2.803679 0.0173964 -1.931152  
 -0.330074 -1.061426 -0.356056 -0.142527 1.0555296 0.3911963 0.639211 0.6169347 -0.174928 0.5212952 0.2576165 -0.931401 0.8330928 0.1536273  
 -0.372388 -0.883228 -0.578208 0.1135003 0.6210344 -0.160557 0.6140334 -0.294722 1.3029012 -2.049291 -1.113813 -0.276684 -1.981971 3.495902  
 -0.294778 -0.205959 -0.169746 0.2250065 0.6088033 -0.72751 0.2895955 -1.650659 0.2140451 1.4910073 0.7012979 1.0159662 3.0282768 0.8495036  
 -0.173662 -0.274804 -0.531462 -0.017641 -0.304435 0.7387889 0.4534711 0.1425863 -1.240246 -2.818243 1.7268374 2.6608127 0.9874708 -0.492319  
 -0.102403 0.6995347 -0.565248 0.4664748 -1.06338 1.7960611 -1.064817 -0.858709 0.0281367 -0.802959 0.3043886 -0.991947 -1.752355 -2.243415  
 0.1023492 0.9036131 -0.328644 0.6083026 -1.788447 1.5305841 -1.083056 0.1216465 0.3538975 1.243145 -2.660189 -0.871952 1.6304743 0.9133453  
 0.1853352 1.167864 0.1834359 0.0441479 -1.751885 0.6362198 -0.408941 2.1110977 -0.593812 0.1516416 1.5483533 -0.989847 0.176 0.2971414  
 0.4380202 2.7624627 0.7257962 -2.68024 -0.573323 -3.45143 -1.861661 -0.968953 -0.038839 0.3666242 -0.446956 1.5026693 -1.22375 -0.057322  
 2.11129 -1.874065 5.9245817 4.309076 -1.257552 -1.672872 0.396552 0.5423309 1.7361525 0.1390657 0.0752437 1.0650887 -0.810331 -0.803772  
 -0.256076 0.2064896 -0.043037 0.119981 -0.218362 -0.109376 0.44414 -0.057797 -0.031533 0.0895605 0.0331169 -0.01612 -0.033545 -0.004877  
 1.8857925 -1.522541 0.3241241 -0.881632 1.6005071 0.7979072 -3.263918 0.4224211 0.198234 -0.667382 -0.231097 0.0977683 0.6288491 0.0537763

DECOMPOSIÇÃO DA INERCIA DOS PERFIS COLUNA

IC

|           |           |           |           |           |           |           |           |           |           |           |           |           |           |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.0066344 | 0.0038448 | 0.0006101 | 0.0007996 | 0.0003007 | 0.0000203 | 2.6735E-7 | 5.8436E-6 | 9.1249E-9 | 2.1249E-7 | 1.0123E-8 | 7.3187E-9 | 8.8769E-9 | 1.3761E-9 |
| 0.0109591 | 0.0021113 | 0.0000619 | 0.0001415 | 0.0002347 | 3.2343E-6 | 0.0000302 | 0.0000144 | 1.3364E-6 | 9.6101E-7 | 1.0001E-7 | 2.4261E-9 | 8.7371E-8 | 2.7759E-8 |
| 0.0111371 | 0.0000267 | 0.0000674 | 0.0000923 | 0.000025  | 3.5825E-7 | 0.0000234 | 9.1542E-6 | 2.6855E-7 | 1.8055E-6 | 2.2042E-6 | 2.9193E-8 | 3.855E-8  | 2.2961E-7 |
| 0.003655  | 2.7033E-6 | 0.0000343 | 3.6997E-6 | 1.4607E-6 | 1.3909E-7 | 0.0000358 | 0.0000248 | 4.6067E-7 | 4.27E-6   | 6.0981E-6 | 2.4559E-8 | 1.4873E-7 | 4.6086E-7 |
| 0.000229  | 0.0007001 | 0.0006153 | 0.0005571 | 0.0000133 | 2.6135E-6 | 0.0000653 | 0.000027  | 1.2709E-6 | 6.2282E-6 | 1.6192E-6 | 2.1419E-7 | 7.7434E-7 | 6.6475E-8 |
| 0.0000488 | 2.9565E-6 | 0.0002485 | 0.0000388 | 5.8585E-6 | 0.000044  | 2.5977E-6 | 3.406E-6  | 0.0000667 | 1.8826E-6 | 4.6612E-9 | 2.1791E-6 | 2.8993E-8 | 2.5015E-7 |
| 0.0001182 | 0.0002008 | 0.0007832 | 0.0000326 | 0.0004039 | 0.0002823 | 0.0000128 | 0.0000196 | 2.5109E-7 | 2.8986E-6 | 5.4763E-7 | 4.0742E-8 | 3.1375E-7 | 6.2223E-8 |
| 0.0000465 | 0.0010539 | 0.0000465 | 0.0000146 | 0.0001314 | 9.3645E-7 | 0.0000144 | 0.0000221 | 0.0000165 | 8.6976E-6 | 3.1717E-6 | 7.3304E-7 | 8.7508E-7 | 2.4989E-7 |
| 7.161E-7  | 0.0003254 | 0.000064  | 0.0000568 | 0.000157  | 0.0000774 | 7.072E-6  | 3.64E-6   | 8.6554E-6 | 3.8397E-9 | 1.0982E-6 | 2.7906E-6 | 9.098E-11 | 7.1535E-7 |
| 0.000165  | 0.0006878 | 0.0000238 | 2.3412E-6 | 0.0000896 | 5.6731E-6 | 0.0000104 | 2.4899E-6 | 1.2344E-7 | 6.3768E-7 | 6.538E-8  | 3.4315E-7 | 2.3586E-7 | 5.0445E-9 |
| 0.0002064 | 0.0002802 | 0.0000617 | 1.4599E-6 | 0.0000305 | 9.396E-7  | 8.0426E-6 | 5.5861E-7 | 6.7329E-6 | 9.6893E-6 | 1.2016E-6 | 2.9774E-8 | 1.2938E-6 | 2.5683E-6 |
| 0.0001298 | 0.0000256 | 5.5333E-6 | 5.7573E-6 | 0.0000294 | 0.0000194 | 1.7887E-6 | 0.0000176 | 1.8236E-7 | 5.1474E-6 | 4.7808E-7 | 4.0287E-7 | 3.0312E-6 | 1.522E-7  |
| 0.0000443 | 0.0000448 | 0.0000514 | 3.4818E-8 | 7.235E-6  | 0.0000196 | 4.3307E-6 | 1.2309E-7 | 6.0234E-6 | 0.0000181 | 2.8517E-6 | 2.7186E-6 | 3.1708E-7 | 5.0289E-8 |
| 0.0000151 | 0.0002944 | 0.0000571 | 0.0000239 | 0.0000866 | 0.0001139 | 0.0000234 | 4.5918E-6 | 3.0404E-9 | 1.4404E-6 | 8.69E-8   | 3.7056E-7 | 9.7934E-7 | 1.0241E-6 |
| 0.0000144 | 0.0004538 | 0.0000184 | 0.0000388 | 0.0002342 | 0.0000791 | 0.0000232 | 8.8114E-8 | 4.5994E-7 | 3.3014E-6 | 6.3466E-6 | 2.7379E-7 | 8.1072E-7 | 1.6232E-7 |
| 0.000044  | 0.0007044 | 5.3404E-6 | 1.9005E-7 | 0.0002088 | 0.0000127 | 3.0697E-6 | 0.0000247 | 1.2035E-6 | 4.5654E-8 | 1.9983E-6 | 3.2792E-7 | 8.7793E-9 | 1.5967E-8 |
| 0.0000004 | 0.0002133 | 0.0000682 | 0.000571  | 0.0000182 | 0.0000304 | 0.0000519 | 4.2355E-6 | 4.197E-9  | 2.1755E-7 | 1.3574E-7 | 6.1606E-7 | 3.4601E-7 | 4.844E-10 |
| 0.0025531 | 0.0008111 | 0.0024908 | 0.0008095 | 0.0000481 | 0.0000392 | 1.2912E-6 | 7.2775E-7 | 4.5997E-6 | 1.7167E-6 | 2.1099E-9 | 1.6975E-7 | 8.321E-8  | 5.2966E-8 |
| 0.0008305 | 0.0002178 | 2.9064E-6 | 0.0000139 | 0.0000321 | 3.7099E-6 | 0.0000358 | 1.8278E-7 | 3.3553E-8 | 1.5745E-7 | 9.0383E-9 | 8.598E-10 | 1.9559E-8 | 4.252E-11 |
| 0.0061383 | 0.0016134 | 0.0000225 | 0.0001021 | 0.0002348 | 0.0000269 | 0.0002635 | 1.3306E-6 | 1.8072E-7 | 1.1915E-6 | 5.998E-8  | 4.3105E-9 | 1.5006E-7 | 7.047E-10 |

INERCIA DOS PERFIS COLUNA

IC

|           |
|-----------|
| 0.0122164 |
| 0.0105804 |
| 0.011386  |
| 0.0037685 |
| 0.0022199 |
| 0.000466  |
| 0.0036576 |
| 0.0013606 |
| 0.0006754 |
| 0.0009885 |
| 0.0006113 |
| 0.000244  |
| 0.0002019 |
| 0.0006128 |
| 0.0008733 |
| 0.0010068 |
| 0.0044331 |
| 0.0067588 |
| 0.0011371 |
| 0.0084044 |

CONTRIBUIÇÃO RELATIVA DOS PERFIS COLUNA

IC

|           |           |           |           |           |           |           |           |           |           |           |           |           |           |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.5430764 | 0.3147276 | 0.0499438 | 0.0654526 | 0.0246162 | 0.0016628 | 0.0000219 | 0.0004783 | 7.4694E-7 | 0.0000174 | 8.2865E-7 | 5.9909E-7 | 7.2664E-7 | 1.1264E-7 |
| 0.8711211 | 0.088469  | 0.0065106 | 0.0112452 | 0.0196555 | 0.0002571 | 0.0023968 | 0.0011448 | 0.0001062 | 0.0000764 | 7.9493E-6 | 1.9294E-7 | 6.945E-6  | 2.2063E-6 |
| 0.9781399 | 0.0023431 | 0.0059166 | 0.0061071 | 0.0021977 | 0.0000315 | 0.0000567 | 0.0000804 | 0.0000236 | 0.0001603 | 0.0001936 | 2.564E-6  | 3.3858E-6 | 0.0000201 |
| 0.9696823 | 0.0007172 | 0.0090918 | 0.0009815 | 0.0003875 | 0.0000369 | 0.0094903 | 0.0065714 | 0.0001222 | 0.0011328 | 0.0016178 | 6.5157E-6 | 0.0000395 | 0.0001223 |
| 0.1031465 | 0.3153624 | 0.2771755 | 0.2509795 | 0.0059962 | 0.0011773 | 0.029437  | 0.0121428 | 0.0005725 | 0.0028057 | 0.0007294 | 0.0000965 | 0.0003488 | 0.0000299 |
| 0.1047486 | 0.006345  | 0.5333662 | 0.0832448 | 0.012573  | 0.0943361 | 0.0055749 | 0.0073076 | 0.1431779 | 0.0040403 | 0.00001   | 0.0046765 | 0.0000622 | 0.0005368 |
| 0.0323221 | 0.5470296 | 0.214113  | 0.0089045 | 0.1104361 | 0.0771901 | 0.0035109 | 0.0053717 | 0.0000686 | 0.0007896 | 0.0001497 | 0.0000111 | 0.0000658 | 0.000017  |
| 0.0341704 | 0.7745553 | 0.0341902 | 0.0107595 | 0.0965565 | 0.0006883 | 0.0106186 | 0.0162413 | 0.0121309 | 0.0063927 | 0.0023312 | 0.0005388 | 0.0006432 | 0.000183  |
| 0.0010603 | 0.481835  | 0.0799664 | 0.0545229 | 0.2324409 | 0.1146565 | 0.0104708 | 0.0053894 | 0.0128152 | 5.6851E-6 | 0.0016261 | 0.0041316 | 1.3471E-7 | 0.0010591 |
| 0.1668822 | 0.6968248 | 0.0240588 | 0.0023685 | 0.0906427 | 0.0057394 | 0.010538  | 0.0025186 | 0.0001249 | 0.0006451 | 0.0000661 | 0.0003472 | 0.0002386 | 5.1034E-6 |
| 0.3376985 | 0.4583555 | 0.1008689 | 0.002388  | 0.0498854 | 0.001537  | 0.0131564 | 0.0009138 | 0.011014  | 0.0158502 | 0.0019657 | 0.0000487 | 0.0021165 | 0.0042014 |
| 0.5320393 | 0.1047248 | 0.0218577 | 0.0235961 | 0.120535  | 0.0793452 | 0.0073307 | 0.0720703 | 0.0007474 | 0.0210962 | 0.0019594 | 0.0016512 | 0.0124231 | 0.0006238 |
| 0.2195037 | 0.2216206 | 0.2546979 | 0.0001724 | 0.0358281 | 0.0972657 | 0.021446  | 0.0006393 | 0.0296284 | 0.0895939 | 0.0141218 | 0.0134627 | 0.0015702 | 0.000249  |
| 0.0246663 | 0.4641211 | 0.0931127 | 0.0389611 | 0.1412736 | 0.1857848 | 0.0382161 | 0.007493  | 4.9615E-6 | 0.0023505 | 0.0001418 | 0.0006047 | 0.0015981 | 0.0016712 |
| 0.0165329 | 0.5196105 | 0.0211194 | 0.0444544 | 0.2681247 | 0.0905279 | 0.0265277 | 0.0001009 | 0.0005266 | 0.0037802 | 0.0072671 | 0.0003135 | 0.0009283 | 0.0001859 |
| 0.0437057 | 0.6996536 | 0.0053045 | 0.0001888 | 0.2074145 | 0.0126103 | 0.003049  | 0.0244978 | 0.0011964 | 0.0000453 | 0.0019848 | 0.0003257 | 8.7203E-6 | 0.0000159 |
| 0.0451967 | 0.7248435 | 0.0153744 | 0.1288133 | 0.0041126 | 0.0687076 | 0.0116987 | 0.0009555 | 9.4673E-7 | 0.0000491 | 0.0000306 | 0.000139  | 0.0000781 | 1.0927E-7 |
| 0.3777419 | 0.1200083 | 0.3665239 | 0.1197741 | 0.007118  | 0.0058065 | 0.000191  | 0.0001077 | 0.0006805 | 2.54E-6   | 3.1217E-7 | 0.0000251 | 0.0000123 | 7.8367E-6 |
| 0.7304263 | 0.1915004 | 0.002556  | 0.0122066 | 0.0282098 | 0.0032627 | 0.0314845 | 0.0001607 | 0.0000295 | 0.0001385 | 7.9487E-6 | 7.5619E-7 | 0.0000172 | 3.7394E-6 |
| 0.7303668 | 0.1919662 | 0.0026732 | 0.0121513 | 0.027943  | 0.0032014 | 0.0313508 | 0.0001583 | 0.0000215 | 0.0001418 | 7.1367E-6 | 5.1289E-7 | 0.0000179 | 8.3844E-8 |

CONTRIBUIÇÃO ABSOLUTADOS PERFIS COLUNA

OPC

|           |           |           |           |           |           |           |           |           |           |           |           |           |           |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.1536811 | 0.2208824 | 0.1140748 | 0.2433277 | 0.1311525 | 0.0192176 | 0.0004322 | 0.0313333 | 0.0000793 | 0.0031759 | 0.0003604 | 0.0006489 | 0.0009294 | 0.0002258 |
| 0.2538573 | 0.0639395 | 0.0153138 | 0.0430509 | 0.1003558 | 0.0030699 | 0.0487448 | 0.0772215 | 0.0116189 | 0.0143632 | 0.0035603 | 0.0002151 | 0.0091475 | 0.0045549 |
| 0.2579804 | 0.0015327 | 0.0125952 | 0.0290902 | 0.1109113 | 0.0003389 | 0.0378563 | 0.049085  | 0.0023349 | 0.0272338 | 0.0784705 | 0.0025884 | 0.0040361 | 0.0375118 |
| 0.0846655 | 0.0001553 | 0.0064073 | 0.0011259 | 0.0006371 | 0.0001316 | 0.0579281 | 0.1328155 | 0.0040052 | 0.0638188 | 0.2170979 | 0.0021775 | 0.0155718 | 0.0756216 |
| 0.0053039 | 0.0402181 | 0.1150399 | 0.1695461 | 0.0058052 | 0.0024726 | 0.1056382 | 0.1445358 | 0.0110493 | 0.0930869 | 0.0576455 | 0.0189908 | 0.061071  | 0.0109079 |
| 0.0011306 | 0.0001698 | 0.0464662 | 0.0119039 | 0.002555  | 0.0415865 | 0.0041994 | 0.0182579 | 0.5800362 | 0.0081373 | 0.0001659 | 0.1932026 | 0.0030355 | 0.0410461 |
| 0.0027385 | 0.1149465 | 0.1464233 | 0.0099114 | 0.1761669 | 0.2671106 | 0.0007594 | 0.1053513 | 0.0021831 | 0.0431737 | 0.0194961 | 0.0056123 | 0.0328489 | 0.01021   |
| 0.0010769 | 0.0005414 | 0.0086973 | 0.0044548 | 0.0572941 | 0.000886  | 0.0233557 | 0.118486  | 0.1434968 | 0.1299944 | 0.1129147 | 0.064993  | 0.0916184 | 0.0408555 |
| 0.0000166 | 0.0186958 | 0.0101005 | 0.0112063 | 0.0664678 | 0.0732638 | 0.0114325 | 0.0195178 | 0.0752526 | 0.0000574 | 0.0390981 | 0.2474122 | 9.5254E-6 | 0.1173809 |
| 0.003821  | 0.0395128 | 0.0044463 | 0.0007124 | 0.039075  | 0.0053672 | 0.0168389 | 0.0133496 | 0.0010732 | 0.0095307 | 0.0023276 | 0.0304251 | 0.0246933 | 0.0008277 |
| 0.0047819 | 0.0160969 | 0.0115287 | 0.0004442 | 0.0132997 | 0.0008869 | 0.0130015 | 0.0029953 | 0.0585374 | 0.1448164 | 0.0427794 | 0.0026398 | 0.135458  | 0.4214332 |
| 0.0030071 | 0.001468  | 0.0009971 | 0.001752  | 0.0123264 | 0.018316  | 0.0028915 | 0.0942903 | 0.0015855 | 0.0769328 | 0.0170199 | 0.0357199 | 0.3173536 | 0.0249737 |
| 0.0010268 | 0.002571  | 0.0096162 | 0.0000106 | 0.0031554 | 0.0185823 | 0.007001  | 0.0006922 | 0.0523691 | 0.2704061 | 0.1015225 | 0.2410394 | 0.0331977 | 0.0082519 |
| 0.0003501 | 0.0163396 | 0.0106684 | 0.0072657 | 0.0377571 | 0.1077121 | 0.0378592 | 0.0246215 | 0.0000264 | 0.0215283 | 0.0030937 | 0.0328548 | 0.1025337 | 0.1680512 |
| 0.0003345 | 0.0260699 | 0.0034485 | 0.0118145 | 0.1021238 | 0.0747979 | 0.0374522 | 0.0004725 | 0.0039988 | 0.0493422 | 0.2259435 | 0.024275  | 0.0848795 | 0.0266345 |
| 0.0010193 | 0.0404666 | 0.0009985 | 0.0000578 | 0.0910712 | 0.0120111 | 0.0049624 | 0.1322472 | 0.0104633 | 0.0006823 | 0.0711394 | 0.0090741 | 0.0009192 | 0.00262   |
| 0.0046412 | 0.1846006 | 0.0127429 | 0.1737752 | 0.0079513 | 0.2881631 | 0.0838381 | 0.0227115 | 0.0000365 | 0.0032515 | 0.0048325 | 0.0546219 | 0.0362264 | 0.0000795 |
| 0.0591396 | 0.0465973 | 0.4656912 | 0.2463491 | 0.009814  | 0.0371286 | 0.0020874 | 0.0039022 | 0.0399906 | 0.0002566 | 0.0000751 | 0.0150506 | 0.0087118 | 0.0086912 |
| 0.0192389 | 0.0125085 | 0.0005434 | 0.0042235 | 0.0139894 | 0.0036099 | 0.0578741 | 0.0009801 | 0.0002917 | 0.0023533 | 0.0003218 | 0.0000762 | 0.0020478 | 6.977E-6  |
| 0.1421887 | 0.0926863 | 0.0042005 | 0.0310779 | 0.1024218 | 0.0054555 | 0.4259471 | 0.0071346 | 0.0015712 | 0.0178084 | 0.0021353 | 0.0003822 | 0.015711  | 0.0001156 |

**B.4 PROGAMA DE AC COM O PROC CORRESP**

```

TITLE2 ' BRASIL: CENSO DEMOGRAFICO DE 1970, 1980 E 1991 (AMOSTRA DE 1,3 E
15%)';
TITLE3 ' LOCALIZACAO DE HOMENS, SEM CONSIDERAR ETAPA DE VIDA, EM SÃO PAULO';
TITLE4 ' LOCALIZACAO, SEGUNDO RENDA, INSTRUCAO E ZONA';
Data hLG2T;
  input local$  iseminst  il_4anos  i5_8anos  i9_1lano  i12_17an  isemdec
    rsemred  decil1  decil2  decil3  decil4  decil5  decil6  decil7  decil8
    decil9  decil10  rsemdec  zurb  zzur;
  label iseminst='i0' il_4anos='i1' i5_8anos='i2' i9_1lano='i3' 2_17an='i4'
    isemdec='in' rsemred='r0' decil1='r1' decil2='r2' decil3='r3'
    decil4='r4' decil5='r5' decil6='r6' decil7='r7' decil8='r8'
    decil9='r9' decil10='r10' rsemdec='rn' zurb='zu' zzur='zr';

cards;
L71 1376 2937 947 924 3731 0 1362 117 0 520 235
    217 943 1119 1075 1738 1945 644 8854
L72 17412 17308 1887 2314 3588 0 2599 102 1510 1105 1193
    3107 2711 7398 7491 6314 7479 1500 31047
L73 12804 17426 895 1101 1398 100 396 1802 3815 2704 3203
    3186 3002 3610 3404 3700 2202 2700 28324
L74 11134 34332 6388 4131 3212 411 306 2760 3676 3906 4892
    4967 5002 7166 8912 7334 4126 6561 50127
L75 139196 773127 300640 165657 84383 740 5779 141192 147193 156616 143004
    141287 143983 124585 125458 119540 88849 123874 1152049
L81 5239 13379 7469 6490 3717 116 16 344 181 556 1284
    1389 3602 6050 7315 7956 7485 232 34593
L82 17729 28753 7948 3899 5447 90 1033 663 5213 2113 5233
    5557 6042 8093 8176 8164 13073 506 55341
L83 17032 21996 3851 1646 1597 28 348 4739 4613 4074 4165
    4987 4723 4404 4743 4556 4268 530 40577
L84 17343 66950 36209 16647 4954 716 479 6611 7812 10535 12001
    12972 17379 22104 21933 18589 10268 2136 133178
L85 113980 679286 787356 387996 95207 9731 3411 209603 218893 246437 238130
    223176 224880 206609 185872 165652 130123 20770 1833664
L91 10344 42139 79224 40694 11066 271 1917 14187 15282 17956 17993
    19530 19049 18426 16976 18554 16804 7065 170980
L92 17367 39459 17612 12457 15109 10 3411 2373 8705 3634 4710
    9659 6489 9840 11184 13906 26550 1554 93086
L93 3330 8830 4685 2711 1706 21 436 1452 1535 1731 1874
    2096 1660 2248 2726 2599 2376 549 20204
L94 5632 21918 45982 22336 6099 162 913 6887 8102 10040 9644
    11469 10200 10509 11233 10610 8319 4202 95749
L95 132670 567576 1053442 482949 119372 3323 18593 195481 209893 247226 249694
    256136 238735 242994 229091 210788 165260 95442 2184566
;
proc corresp data=HLG2T out=SHLG2T ALL profile=row;
  VAR iseminst il_4anos i5_8anos i9_1lano i12_17an isemdec rsemred
    decil1 decil2 decil3 decil4 decil5 decil6 decil7 decil8 decil9
    decil10 rsemdec zurb zzur;
  ID LOCAL;
RUN;

DATA GHLG2T;
  SET SHLG2T;
  X=DIM1;
  Y=DIM2;
  XSYS='2';
  YSYS='2';

```

```
TEXT= LOCAL;
SIZE=1.5;
LABEL X='DIMENSAO 1'
      Y='DIMENSAO 2';
KEEP X Y TEXT XSYS YSYS SIZE;

proc print data=gHLG2T;
RUN;

FILENAME FIG 'C:\USERS\ADMUR\tese3\GHLG2TS.GRF';
GOPTIONS DEVICE=CGMWPWA GSFNAME=FIG GSFMODE=REPLACE;

PROC GPLOT DATA=GHLG2T;
  SYMBOL VALUE=NONE;
  AXIS1 ORDER=-2 TO 8 BY 1;
  PLOT Y*X=1 / ANNOTATE=GHLG2T
           VAXIS=AXIS1 HAXIS=AXIS1
           VREF=0 HREF=0
           FRAME;
RUN;

GOPTIONS DEVICE=WIN;
QUIT;

FILENAME FIG 'C:\USERS\ADMUR\tese3\GHLG2TZ.GRF';
GOPTIONS DEVICE=CGMWPWA GSFNAME=FIG GSFMODE=REPLACE;

PROC GPLOT DATA=GHLG2T;
  SYMBOL VALUE=NONE;
  AXIS1 ORDER=-0.5 TO 1 BY 0.5;
  PLOT Y*X=1 / ANNOTATE=GHLG2T
           VAXIS=AXIS1 HAXIS=AXIS1
           VREF=0 HREF=0
           FRAME;
RUN;

GOPTIONS DEVICE=WIN;
QUIT;
```

**B.5 SAÍDA DO PROC CORRESP/SAS**

The SAS System  
The Correspondence Analysis Procedure

Contingency Table

|     | i0     | i1      | i2      | i3      | i4     | in    | r0    |
|-----|--------|---------|---------|---------|--------|-------|-------|
| L71 | 1376   | 2937    | 947     | 924     | 3731   | 0     | 1362  |
| L72 | 17412  | 17308   | 1887    | 2314    | 3588   | 0     | 2599  |
| L73 | 12804  | 17426   | 895     | 1101    | 1398   | 100   | 396   |
| L74 | 11134  | 34332   | 6388    | 4131    | 3212   | 411   | 306   |
| L75 | 139196 | 773127  | 300640  | 165657  | 84383  | 740   | 5779  |
| L81 | 5239   | 13379   | 7469    | 6490    | 3717   | 116   | 16    |
| L82 | 17729  | 28753   | 7948    | 3899    | 5447   | 90    | 1033  |
| L83 | 17032  | 21996   | 3851    | 1646    | 1597   | 28    | 348   |
| L84 | 17343  | 66950   | 36209   | 16647   | 4954   | 716   | 479   |
| L85 | 113980 | 679286  | 787356  | 387996  | 95207  | 9731  | 3411  |
| L91 | 10344  | 42139   | 79224   | 40694   | 11066  | 271   | 1917  |
| L92 | 17367  | 39459   | 17612   | 12457   | 15109  | 10    | 3411  |
| L93 | 3330   | 8830    | 4685    | 2711    | 1706   | 21    | 436   |
| L94 | 5632   | 21918   | 45982   | 22336   | 6099   | 162   | 913   |
| L95 | 132670 | 567576  | 1053442 | 482949  | 119372 | 3323  | 18593 |
| Sum | 522588 | 2335416 | 2354535 | 1151952 | 360586 | 15719 | 40999 |

|     | r1     | r2     | r3     | r4     | r5     | r6     | r7     |
|-----|--------|--------|--------|--------|--------|--------|--------|
| L71 | 117    | 0      | 520    | 235    | 217    | 943    | 1119   |
| L72 | 102    | 1510   | 1105   | 1193   | 3107   | 2711   | 7398   |
| L73 | 1802   | 3815   | 2704   | 3203   | 3186   | 3002   | 3610   |
| L74 | 2760   | 3676   | 3906   | 4892   | 4967   | 5002   | 7166   |
| L75 | 141192 | 147193 | 156616 | 143004 | 141287 | 143983 | 124585 |
| L81 | 344    | 181    | 556    | 1284   | 1389   | 3602   | 6050   |
| L82 | 663    | 5213   | 2113   | 5233   | 5557   | 6042   | 8093   |
| L83 | 4739   | 4613   | 4074   | 4165   | 4987   | 4723   | 4404   |
| L84 | 6611   | 7812   | 10535  | 12001  | 12972  | 17379  | 22104  |
| L85 | 209603 | 218893 | 246437 | 238130 | 223176 | 224880 | 206609 |
| L91 | 14187  | 15282  | 17956  | 17993  | 19530  | 19049  | 18426  |
| L92 | 2373   | 8705   | 3634   | 4710   | 9659   | 6489   | 9840   |
| L93 | 1452   | 1535   | 1731   | 1874   | 2096   | 1660   | 2248   |
| L94 | 6887   | 8102   | 10040  | 9644   | 11469  | 10200  | 10509  |
| L95 | 195481 | 209893 | 247226 | 249694 | 256136 | 238735 | 242994 |
| Sum | 588313 | 636423 | 709153 | 697255 | 699735 | 688400 | 675155 |

|     | r8     | r9     | r10    | rn     | zu      | zr     | Sum      |
|-----|--------|--------|--------|--------|---------|--------|----------|
| L71 | 1075   | 1738   | 1945   | 644    | 8854    | 1061   | 29745    |
| L72 | 7491   | 6314   | 7479   | 1500   | 31047   | 11462  | 127527   |
| L73 | 3404   | 3700   | 2202   | 2700   | 28324   | 5400   | 101172   |
| L74 | 8912   | 7334   | 4126   | 6561   | 50127   | 9481   | 178824   |
| L75 | 125458 | 119540 | 88849  | 123874 | 1152049 | 311694 | 4388846  |
| L81 | 7315   | 7956   | 7485   | 232    | 34593   | 1817   | 109230   |
| L82 | 8176   | 8164   | 13073  | 506    | 55341   | 8525   | 191598   |
| L83 | 4743   | 4556   | 4268   | 530    | 40577   | 5573   | 138450   |
| L84 | 21933  | 18589  | 10268  | 2136   | 133178  | 9641   | 428457   |
| L85 | 185872 | 165652 | 130123 | 20770  | 1833664 | 239892 | 6220668  |
| L91 | 16976  | 18554  | 16804  | 7065   | 170980  | 12760  | 551217   |
| L92 | 11184  | 13906  | 26550  | 1554   | 93086   | 8929   | 306044   |
| L93 | 2726   | 2599   | 2376   | 549    | 20204   | 1080   | 63849    |
| L94 | 11233  | 10610  | 8319   | 4202   | 95749   | 6379   | 306385   |
| L95 | 229091 | 210788 | 165260 | 95442  | 2184566 | 174766 | 7077997  |
| Sum | 645589 | 600000 | 489127 | 268265 | 5932339 | 908460 | 20220009 |

Row Profiles

|     | i0       | i1       | i2       | i3       | i4       | in       | r0       |
|-----|----------|----------|----------|----------|----------|----------|----------|
| L71 | 0.046260 | 0.098739 | 0.031837 | 0.031064 | 0.125433 | 0.000000 | 0.045789 |
| L72 | 0.136536 | 0.135720 | 0.014797 | 0.018145 | 0.028135 | 0.000000 | 0.020380 |
| L73 | 0.126557 | 0.172241 | 0.008846 | 0.010882 | 0.013818 | 0.000988 | 0.003914 |
| L74 | 0.062262 | 0.191988 | 0.035722 | 0.023101 | 0.017962 | 0.002298 | 0.001711 |
| L75 | 0.031716 | 0.176157 | 0.068501 | 0.037745 | 0.019227 | 0.000169 | 0.001317 |
| L81 | 0.047963 | 0.122485 | 0.068379 | 0.059416 | 0.034029 | 0.001062 | 0.000146 |
| L82 | 0.092532 | 0.150069 | 0.041483 | 0.020350 | 0.028429 | 0.000470 | 0.005391 |
| L83 | 0.123019 | 0.158873 | 0.027815 | 0.011889 | 0.011535 | 0.000202 | 0.002514 |
| L84 | 0.040478 | 0.156258 | 0.084510 | 0.038853 | 0.011562 | 0.001671 | 0.001118 |
| L85 | 0.018323 | 0.109199 | 0.126571 | 0.062372 | 0.015305 | 0.001564 | 0.000548 |
| L91 | 0.018766 | 0.076447 | 0.143726 | 0.073826 | 0.020076 | 0.000492 | 0.003478 |
| L92 | 0.056747 | 0.128932 | 0.057547 | 0.040703 | 0.049369 | 0.000033 | 0.011145 |
| L93 | 0.052154 | 0.138295 | 0.073376 | 0.042460 | 0.026719 | 0.000329 | 0.006829 |
| L94 | 0.018382 | 0.071537 | 0.150079 | 0.072902 | 0.019906 | 0.000529 | 0.002980 |
| L95 | 0.018744 | 0.080189 | 0.148833 | 0.068232 | 0.016865 | 0.000469 | 0.002627 |

|     | r1       | r2       | r3       | r4       | r5       | r6       | r7       |
|-----|----------|----------|----------|----------|----------|----------|----------|
| L71 | 0.003933 | 0.000000 | 0.017482 | 0.007900 | 0.007295 | 0.031703 | 0.037620 |
| L72 | 0.000800 | 0.011841 | 0.008665 | 0.009355 | 0.024363 | 0.021258 | 0.058011 |
| L73 | 0.017811 | 0.037708 | 0.026727 | 0.031659 | 0.031491 | 0.029672 | 0.035682 |
| L74 | 0.015434 | 0.020557 | 0.021843 | 0.027357 | 0.027776 | 0.027972 | 0.040073 |
| L75 | 0.032171 | 0.033538 | 0.035685 | 0.032584 | 0.032192 | 0.032907 | 0.028387 |
| L81 | 0.003149 | 0.001657 | 0.005090 | 0.011755 | 0.012716 | 0.032976 | 0.055388 |
| L82 | 0.003460 | 0.027208 | 0.011028 | 0.027312 | 0.029003 | 0.031535 | 0.042239 |
| L83 | 0.034229 | 0.033319 | 0.029426 | 0.030083 | 0.036020 | 0.034113 | 0.031809 |
| L84 | 0.015430 | 0.018233 | 0.024588 | 0.028010 | 0.030276 | 0.040562 | 0.051590 |
| L85 | 0.033695 | 0.035188 | 0.039616 | 0.038280 | 0.035877 | 0.036150 | 0.033213 |
| L91 | 0.025738 | 0.027724 | 0.032575 | 0.032642 | 0.035431 | 0.034558 | 0.033428 |
| L92 | 0.007754 | 0.028444 | 0.011874 | 0.015390 | 0.031561 | 0.021203 | 0.032152 |
| L93 | 0.022741 | 0.024041 | 0.027111 | 0.029350 | 0.032827 | 0.025999 | 0.035208 |
| L94 | 0.022478 | 0.026444 | 0.032769 | 0.031477 | 0.037433 | 0.033291 | 0.034300 |
| L95 | 0.027618 | 0.029654 | 0.034929 | 0.035277 | 0.036188 | 0.033729 | 0.034331 |

|     | r8       | r9       | r10      | rn       | zu       | zr       |
|-----|----------|----------|----------|----------|----------|----------|
| L71 | 0.036141 | 0.058430 | 0.065389 | 0.021651 | 0.297663 | 0.035670 |
| L72 | 0.058741 | 0.049511 | 0.058646 | 0.011762 | 0.243454 | 0.089879 |
| L73 | 0.033646 | 0.036571 | 0.021765 | 0.026687 | 0.279959 | 0.053374 |
| L74 | 0.049837 | 0.041012 | 0.023073 | 0.036690 | 0.280315 | 0.053019 |
| L75 | 0.028586 | 0.027237 | 0.020244 | 0.028225 | 0.262495 | 0.071020 |
| L81 | 0.066969 | 0.072837 | 0.068525 | 0.002124 | 0.316699 | 0.016635 |
| L82 | 0.042673 | 0.042610 | 0.068231 | 0.002641 | 0.288839 | 0.044494 |
| L83 | 0.034258 | 0.032907 | 0.030827 | 0.003828 | 0.293081 | 0.040253 |
| L84 | 0.051191 | 0.043386 | 0.023965 | 0.004985 | 0.310832 | 0.022502 |
| L85 | 0.029880 | 0.026629 | 0.020918 | 0.003339 | 0.294770 | 0.038564 |
| L91 | 0.030797 | 0.033660 | 0.030485 | 0.012817 | 0.310186 | 0.023149 |
| L92 | 0.036544 | 0.045438 | 0.086752 | 0.005078 | 0.304159 | 0.029176 |
| L93 | 0.042694 | 0.040705 | 0.037213 | 0.008598 | 0.316434 | 0.016915 |
| L94 | 0.036663 | 0.034630 | 0.027152 | 0.013715 | 0.312512 | 0.020820 |
| L95 | 0.032367 | 0.029781 | 0.023348 | 0.013484 | 0.308642 | 0.024691 |

Column Profiles

|     | i0       | i1       | i2       | i3       | i4       | in       | r0       |
|-----|----------|----------|----------|----------|----------|----------|----------|
| L71 | 0.002633 | 0.001258 | 0.000402 | 0.000802 | 0.010347 | 0.000000 | 0.033220 |
| L72 | 0.033319 | 0.007411 | 0.000801 | 0.002009 | 0.009950 | 0.000000 | 0.063392 |
| L73 | 0.024501 | 0.007462 | 0.000380 | 0.000956 | 0.003877 | 0.006362 | 0.009659 |
| L74 | 0.021306 | 0.014701 | 0.002713 | 0.003586 | 0.008908 | 0.026147 | 0.007464 |
| L75 | 0.266359 | 0.331045 | 0.127686 | 0.143805 | 0.234016 | 0.047077 | 0.140955 |
| L81 | 0.010025 | 0.005729 | 0.003172 | 0.005634 | 0.010308 | 0.007380 | 0.000390 |
| L82 | 0.033925 | 0.012312 | 0.003376 | 0.003385 | 0.015106 | 0.005726 | 0.025196 |
| L83 | 0.032592 | 0.009418 | 0.001636 | 0.001429 | 0.004429 | 0.001781 | 0.008488 |
| L84 | 0.033187 | 0.028667 | 0.015378 | 0.014451 | 0.013739 | 0.045550 | 0.011683 |
| L85 | 0.218107 | 0.290863 | 0.334400 | 0.336816 | 0.264034 | 0.619060 | 0.083197 |
| L91 | 0.019794 | 0.018043 | 0.033647 | 0.035326 | 0.030689 | 0.017240 | 0.046757 |
| L92 | 0.033233 | 0.016896 | 0.007480 | 0.010814 | 0.041901 | 0.000636 | 0.083197 |
| L93 | 0.006372 | 0.003781 | 0.001990 | 0.002353 | 0.004731 | 0.001336 | 0.010634 |
| L94 | 0.010777 | 0.009385 | 0.019529 | 0.019390 | 0.016914 | 0.010306 | 0.022269 |
| L95 | 0.253871 | 0.243030 | 0.447410 | 0.419244 | 0.331050 | 0.211400 | 0.453499 |

|     | r1       | r2       | r3       | r4       | r5       | r6       | r7       |
|-----|----------|----------|----------|----------|----------|----------|----------|
| L71 | 0.000199 | 0.000000 | 0.000733 | 0.000337 | 0.000310 | 0.001370 | 0.001657 |
| L72 | 0.000173 | 0.002373 | 0.001558 | 0.001711 | 0.004440 | 0.003938 | 0.010957 |
| L73 | 0.003063 | 0.005994 | 0.003813 | 0.004594 | 0.004553 | 0.004361 | 0.005347 |
| L74 | 0.004691 | 0.005776 | 0.005508 | 0.007016 | 0.007098 | 0.007266 | 0.010614 |
| L75 | 0.239995 | 0.231282 | 0.220949 | 0.205096 | 0.201915 | 0.209156 | 0.184528 |
| L81 | 0.000585 | 0.000284 | 0.000784 | 0.001842 | 0.001985 | 0.005232 | 0.008961 |
| L82 | 0.001127 | 0.008191 | 0.002980 | 0.007505 | 0.007942 | 0.008777 | 0.011987 |
| L83 | 0.008055 | 0.007248 | 0.005745 | 0.005973 | 0.007127 | 0.006861 | 0.006523 |
| L84 | 0.011237 | 0.012275 | 0.014856 | 0.017212 | 0.018538 | 0.025245 | 0.032739 |
| L85 | 0.356278 | 0.343943 | 0.347509 | 0.341525 | 0.318944 | 0.326671 | 0.306017 |
| L91 | 0.024115 | 0.024012 | 0.025320 | 0.025805 | 0.027911 | 0.027671 | 0.027292 |
| L92 | 0.004034 | 0.013678 | 0.005124 | 0.006755 | 0.013804 | 0.009426 | 0.014574 |
| L93 | 0.002468 | 0.002412 | 0.002441 | 0.002688 | 0.002995 | 0.002411 | 0.003330 |
| L94 | 0.011706 | 0.012731 | 0.014158 | 0.013831 | 0.016390 | 0.014817 | 0.015565 |
| L95 | 0.332274 | 0.329801 | 0.348622 | 0.358110 | 0.366047 | 0.346797 | 0.359908 |

|     | r8       | r9       | r10      | rn       | zu       | zr       |
|-----|----------|----------|----------|----------|----------|----------|
| L71 | 0.001665 | 0.002897 | 0.003976 | 0.002401 | 0.001492 | 0.001312 |
| L72 | 0.011603 | 0.010523 | 0.015291 | 0.005591 | 0.005234 | 0.014178 |
| L73 | 0.005273 | 0.006167 | 0.004502 | 0.010065 | 0.004775 | 0.006679 |
| L74 | 0.013804 | 0.012223 | 0.008435 | 0.024457 | 0.008450 | 0.011727 |
| L75 | 0.194331 | 0.199233 | 0.181648 | 0.461760 | 0.194198 | 0.385540 |
| L81 | 0.011331 | 0.013260 | 0.015303 | 0.000865 | 0.005831 | 0.002247 |
| L82 | 0.012664 | 0.013607 | 0.026727 | 0.001886 | 0.009329 | 0.010545 |
| L83 | 0.007347 | 0.007593 | 0.008726 | 0.001976 | 0.006840 | 0.006893 |
| L84 | 0.033974 | 0.030982 | 0.020993 | 0.007962 | 0.022449 | 0.011925 |
| L85 | 0.287911 | 0.276087 | 0.266031 | 0.077423 | 0.309096 | 0.296727 |
| L91 | 0.026295 | 0.030923 | 0.034355 | 0.026336 | 0.028822 | 0.015783 |
| L92 | 0.017324 | 0.023177 | 0.054280 | 0.005793 | 0.015691 | 0.011044 |
| L93 | 0.004223 | 0.004332 | 0.004858 | 0.002046 | 0.003406 | 0.001336 |
| L94 | 0.017400 | 0.017683 | 0.017008 | 0.015664 | 0.016140 | 0.007890 |
| L95 | 0.354856 | 0.351313 | 0.337867 | 0.355775 | 0.368247 | 0.216171 |

Inertia and Chi-Square Decomposition

| Singular Values | Principal Inertias | Chi-Squares | Percents                   | 12    | 24 | 36 | 48 | 60 |
|-----------------|--------------------|-------------|----------------------------|-------|----|----|----|----|
| 0.20777         | 0.04317            | 872901      | 58.65%                     | ***** |    |    |    |    |
| 0.13193         | 0.01741            | 351964      | 23.65%                     | ***** |    |    |    |    |
| 0.07313         | 0.00535            | 108147      | 7.27%                      | ***   |    |    |    |    |
| 0.05732         | 0.00329            | 66444.7     | 4.46%                      | **    |    |    |    |    |
| 0.04788         | 0.00229            | 46362.9     | 3.12%                      | *     |    |    |    |    |
| 0.03251         | 0.00106            | 21372.4     | 1.44%                      | *     |    |    |    |    |
| 0.02487         | 0.00062            | 12507.8     | 0.84%                      |       |    |    |    |    |
| 0.01366         | 0.00019            | 3770.96     | 0.25%                      |       |    |    |    |    |
| 0.01072         | 0.00012            | 2325.67     | 0.16%                      |       |    |    |    |    |
| 0.00818         | 0.00007            | 1352.87     | 0.09%                      |       |    |    |    |    |
| 0.00530         | 0.00003            | 567.967     | 0.04%                      |       |    |    |    |    |
| 0.00336         | 0.00001            | 228.055     | 0.02%                      |       |    |    |    |    |
| 0.00309         | 0.00001            | 193.129     | 0.01%                      |       |    |    |    |    |
| 0.00247         | 0.00001            | 123.226     | 0.01%                      |       |    |    |    |    |
|                 | 0.07360            | 1488263     | (Degrees of Freedom = 266) |       |    |    |    |    |

Row Coordinates (COORDENADAS PRINCIPAIS)

|     | Dim1      | Dim2      |
|-----|-----------|-----------|
| L71 | 0.354496  | 0.824205  |
| L72 | 0.678312  | 0.659366  |
| L73 | 0.610488  | 0.255334  |
| L74 | 0.454618  | 0.072360  |
| L75 | 0.296513  | -0.130912 |
| L81 | 0.125058  | 0.468614  |
| L82 | 0.406989  | 0.458436  |
| L83 | 0.472572  | 0.298989  |
| L84 | 0.116754  | 0.136215  |
| L85 | -0.084193 | -0.045318 |
| L91 | -0.164771 | 0.081095  |
| L92 | 0.244024  | 0.507820  |
| L93 | 0.147970  | 0.253272  |
| L94 | -0.185236 | 0.085318  |
| L95 | -0.164080 | 0.032195  |

Summary Statistics for the Row Points

|     | Quality  | Mass     | Inertia  |
|-----|----------|----------|----------|
| L71 | 0.424677 | 0.001471 | 0.037884 |
| L72 | 0.872718 | 0.006307 | 0.087863 |
| L73 | 0.743210 | 0.005004 | 0.040053 |
| L74 | 0.808417 | 0.008844 | 0.031497 |
| L75 | 0.980082 | 0.217055 | 0.316108 |
| L81 | 0.613292 | 0.005402 | 0.028152 |
| L82 | 0.918121 | 0.009476 | 0.052695 |
| L83 | 0.626405 | 0.006847 | 0.046442 |
| L84 | 0.319669 | 0.021190 | 0.028986 |
| L85 | 0.541518 | 0.307649 | 0.070566 |
| L91 | 0.856814 | 0.027261 | 0.014579 |
| L92 | 0.800250 | 0.015136 | 0.081569 |
| L93 | 0.809634 | 0.003158 | 0.004559 |
| L94 | 0.846063 | 0.015153 | 0.010120 |
| L95 | 0.892860 | 0.350049 | 0.148925 |

Partial Contributions to Inertia for the Row Points (CONTRIBUIÇÃO ABSOLUTA)

|     | Dim1     | Dim2     |
|-----|----------|----------|
| L71 | 0.004282 | 0.057410 |
| L72 | 0.067220 | 0.157528 |
| L73 | 0.043197 | 0.018740 |
| L74 | 0.042340 | 0.002660 |
| L75 | 0.442050 | 0.213701 |
| L81 | 0.001957 | 0.068151 |
| L82 | 0.036357 | 0.114406 |
| L83 | 0.035421 | 0.035164 |
| L84 | 0.006691 | 0.022587 |
| L85 | 0.050516 | 0.036298 |
| L91 | 0.017144 | 0.010299 |
| L92 | 0.020878 | 0.224236 |
| L93 | 0.001602 | 0.011637 |
| L94 | 0.012043 | 0.006337 |
| L95 | 0.218303 | 0.020845 |

Indices of the Coordinates that Contribute Most to Inertia for the Row Points

|     | Dim1 | Dim2 | Best |
|-----|------|------|------|
| L71 | 0    | 2    | 2    |
| L72 | 2    | 2    | 2    |
| L73 | 1    | 0    | 1    |
| L74 | 0    | 0    | 1    |
| L75 | 1    | 1    | 1    |
| L81 | 0    | 2    | 2    |
| L82 | 0    | 2    | 2    |
| L83 | 0    | 0    | 1    |
| L84 | 0    | 0    | 2    |
| L85 | 1    | 0    | 1    |
| L91 | 0    | 0    | 1    |
| L92 | 0    | 2    | 2    |
| L93 | 0    | 0    | 2    |
| L94 | 0    | 0    | 1    |
| L95 | 1    | 0    | 1    |

Squared Cosines for the Row Points (CONTRIBUIÇÃO RELATIVA)

|     | Dim1     | Dim2     |
|-----|----------|----------|
| L71 | 0.066297 | 0.358380 |
| L72 | 0.448717 | 0.424001 |
| L73 | 0.632557 | 0.110653 |
| L74 | 0.788442 | 0.019975 |
| L75 | 0.820203 | 0.159879 |
| L81 | 0.040774 | 0.572518 |
| L82 | 0.404673 | 0.513448 |
| L83 | 0.447340 | 0.179065 |
| L84 | 0.135386 | 0.184282 |
| L85 | 0.419869 | 0.121649 |
| L91 | 0.689738 | 0.167076 |
| L92 | 0.150122 | 0.650128 |
| L93 | 0.206029 | 0.603606 |
| L94 | 0.697988 | 0.148076 |
| L95 | 0.859758 | 0.033102 |

Column Coordinates ( COORDENADAS PADRONIZADAS)

|     | Dim1     | Dim2     |
|-----|----------|----------|
| i0  | 2.43849  | 2.92342  |
| i1  | 1.48253  | -0.74403 |
| i2  | -1.48844 | -0.11473 |
| i3  | -1.21906 | -0.05221 |
| i4  | 0.54536  | 1.50175  |
| in  | -1.20596 | -0.46742 |
| r0  | 1.16215  | 7.52925  |
| r1  | -0.19239 | -1.44249 |
| r2  | -0.02296 | -0.77071 |
| r3  | -0.33007 | -1.06143 |
| r4  | -0.37239 | -0.68323 |
| r5  | -0.29478 | -0.20596 |
| r6  | -0.17366 | -0.27480 |
| r7  | -0.10240 | 0.69953  |
| r8  | 0.10235  | 0.90361  |
| r9  | 0.18534  | 1.16779  |
| r10 | 0.43802  | 2.76246  |
| rn  | 2.11129  | -1.87408 |
| zu  | -0.25608 | 0.20649  |
| zr  | 1.88579  | -1.52254 |

Summary Statistics for the Column Points

|     | Quality  | Mass     | Inertia  |
|-----|----------|----------|----------|
| i0  | 0.857804 | 0.025845 | 0.165976 |
| i1  | 0.959590 | 0.115500 | 0.170922 |
| i2  | 0.980483 | 0.116446 | 0.154693 |
| i3  | 0.970399 | 0.056971 | 0.051211 |
| i4  | 0.418509 | 0.017833 | 0.030160 |
| in  | 0.111094 | 0.000777 | 0.006331 |
| r0  | 0.579352 | 0.002028 | 0.049694 |
| r1  | 0.808726 | 0.029096 | 0.018485 |
| r2  | 0.482895 | 0.031475 | 0.009176 |
| r3  | 0.862707 | 0.035072 | 0.013429 |
| r4  | 0.796054 | 0.034483 | 0.008305 |
| r5  | 0.636764 | 0.034606 | 0.003315 |
| r6  | 0.441124 | 0.034045 | 0.002744 |
| r7  | 0.488787 | 0.033390 | 0.008326 |
| r8  | 0.536143 | 0.031928 | 0.011865 |
| r9  | 0.743359 | 0.029674 | 0.013678 |
| r10 | 0.770040 | 0.024190 | 0.060229 |
| rn  | 0.497750 | 0.013267 | 0.091827 |
| zu  | 0.921927 | 0.293390 | 0.015449 |
| zr  | 0.922333 | 0.039983 | 0.114185 |

Partial Contributions to Inertia for the Column Points (CONTRIBUIÇÃO ABSOLUTA)

|     | Dim1     | Dim2     |
|-----|----------|----------|
| i0  | 0.153681 | 0.220882 |
| i1  | 0.253857 | 0.063939 |
| i2  | 0.257980 | 0.001533 |
| i3  | 0.084666 | 0.000155 |
| i4  | 0.005304 | 0.040218 |
| in  | 0.001131 | 0.000170 |
| r0  | 0.002739 | 0.114946 |
| r1  | 0.001077 | 0.060541 |
| r2  | 0.000017 | 0.018696 |
| r3  | 0.003821 | 0.039513 |
| r4  | 0.004782 | 0.016097 |
| r5  | 0.003007 | 0.001468 |
| r6  | 0.001027 | 0.002571 |
| r7  | 0.000350 | 0.016340 |
| r8  | 0.000334 | 0.026070 |
| r9  | 0.001019 | 0.040467 |
| r10 | 0.004641 | 0.184601 |
| rn  | 0.059140 | 0.046597 |
| zu  | 0.019239 | 0.012510 |
| zr  | 0.142189 | 0.092686 |

Indices of the Coordinates that Contribute Most to Inertia for the Column Points

|     | Dim1 | Dim2 | Best |
|-----|------|------|------|
| i0  | 2    | 2    | 2    |
| i1  | 1    | 1    | 1    |
| i2  | 1    | 0    | 1    |
| i3  | 0    | 0    | 1    |
| i4  | 0    | 0    | 2    |
| in  | 0    | 0    | 1    |
| r0  | 0    | 2    | 2    |
| r1  | 0    | 2    | 2    |
| r2  | 0    | 0    | 2    |
| r3  | 0    | 0    | 2    |
| r4  | 0    | 0    | 2    |
| r5  | 0    | 0    | 1    |
| r6  | 0    | 0    | 2    |
| r7  | 0    | 0    | 2    |
| r8  | 0    | 0    | 2    |
| r9  | 0    | 2    | 2    |
| r10 | 0    | 2    | 2    |
| rn  | 0    | 1    | 1    |
| ru  | 0    | 0    | 1    |
| rr  | 1    | 1    | 1    |

Squared Cosines for the Column Points (CONTRIBUIÇÃO RELATIVA)

|     | Dim1     | Dim2     |
|-----|----------|----------|
| i0  | 0.543076 | 0.314728 |
| i1  | 0.871131 | 0.088469 |
| i2  | 0.978140 | 0.002343 |
| i3  | 0.969682 | 0.000717 |
| i4  | 0.103147 | 0.315362 |
| in  | 0.104749 | 0.006345 |
| r0  | 0.032322 | 0.547030 |
| r1  | 0.034170 | 0.774555 |
| r2  | 0.001060 | 0.481835 |
| r3  | 0.166882 | 0.695825 |
| r4  | 0.337699 | 0.458355 |
| r5  | 0.532039 | 0.104725 |
| r6  | 0.219504 | 0.221621 |
| r7  | 0.024666 | 0.464121 |
| r8  | 0.016533 | 0.519611 |
| r9  | 0.043706 | 0.699654 |
| r10 | 0.045197 | 0.724844 |
| rn  | 0.377742 | 0.120009 |
| ru  | 0.730426 | 0.191500 |
| rr  | 0.730367 | 0.191966 |

**B.6 PROGRAMA DO SAS (PROC CORRESP) PARA OS PONTOS SUPLEMENTARES**

TITLE2 ' BRASIL: CENSO DEMOGRÁFICO DE 1970, 1980 E 1991 (AMOSTRA DE 1, 3 E 15%)';  
 TITLE3 ' LOCALIZAÇÃO DOMÉSTICA SEM CONSIDERAR AS ETAPAS DE VIDA, DE HOMENS SEM COMPANHERAS EM SÃO PAULO';  
 TITLE4 ' LOCALIZAÇÃO, SEGUNDO RENDA, INSTRUÇÃO E ZONA';  
 TITLE5 ' OBS: FAZENDO ESTAS LOCALIZAÇÕES COMO GRUPOS E AS DENTRO DAS ETAPAS COMO PONTOS SUPLEMENTARES';

data HLS2T;

input locals

10 i1 i2 i3 i4 in r0 r1 r2 r3 r4 r5 r6 r7 r8 r9 r10 rn zu zr;

cards;

|      |        |        |        |        |        |         |        |        |        |        |        |        |        |
|------|--------|--------|--------|--------|--------|---------|--------|--------|--------|--------|--------|--------|--------|
| L711 | 752    | 1822   | 947    | 924    | 3413   | 0       | 1362   | 117    | 0      | 420    | 0      | 0      | 943    |
|      | 626    | 954    | 1638   | 1271   | 527    | 7100    | 758    |        |        |        |        |        |        |
| L721 | 503    | 697    | 0      | 0      | 218    | 0       | 0      | 0      | 0      | 100    | 235    | 0      | 0      |
|      | 392    | 0      | 0      | 574    | 117    | 1115    | 303    |        |        |        |        |        |        |
| L731 | 121    | 418    | 0      | 0      | 100    | 0       | 0      | 0      | 0      | 0      | 0      | 217    | 0      |
|      | 101    | 121    | 100    | 100    | 0      | 639     | 0      |        |        |        |        |        |        |
| L71  | 1376   | 2937   | 947    | 924    | 3731   | 0       | 1362   | 117    | 0      | 520    | 235    | 217    | 943    |
|      | 1119   | 1075   | 1738   | 1945   | 644    | 8854    | 1061   |        |        |        |        |        |        |
| L712 | 5200   | 6194   | 1289   | 1105   | 1489   | 0       | 392    | 0      | 308    | 305    | 398    | 495    | 707    |
|      | 3703   | 3288   | 2998   | 2485   | 198    | 11759   | 3518   |        |        |        |        |        |        |
| L722 | 5398   | 7208   | 498    | 1108   | 1800   | 0       | 198    | 0      | 303    | 598    | 199    | 1399   | 505    |
|      | 2303   | 3099   | 2311   | 4196   | 901    | 12185   | 3827   |        |        |        |        |        |        |
| L732 | 6814   | 3906   | 100    | 101    | 299    | 0       | 2009   | 102    | 999    | 202    | 596    | 1213   | 1499   |
|      | 1392   | 1104   | 1005   | 798    | 401    | 7103    | 4117   |        |        |        |        |        |        |
| L72  | 17412  | 17308  | 1887   | 2314   | 3588   | 0       | 2599   | 102    | 1510   | 1105   | 1193   | 3107   | 2711   |
|      | 7398   | 7491   | 6314   | 7479   | 1500   | 31047   | 11462  |        |        |        |        |        |        |
| L713 | 401    | 1205   | 100    | 0      | 297    | 0       | 0      | 403    | 405    | 99     | 199    | 297    | 0      |
|      | 302    | 100    | 198    | 0      | 0      | 1700    | 303    |        |        |        |        |        |        |
| L723 | 2802   | 6021   | 598    | 801    | 700    | 0       | 99     | 503    | 1213   | 1204   | 1501   | 598    | 1003   |
|      | 1202   | 799    | 1404   | 500    | 896    | 9409    | 1513   |        |        |        |        |        |        |
| L733 | 9601   | 10200  | 197    | 300    | 401    | 100     | 297    | 896    | 2197   | 1401   | 1503   | 2291   | 1999   |
|      | 2106   | 2505   | 2098   | 1702   | 1804   | 17215   | 3584   |        |        |        |        |        |        |
| L73  | 12804  | 17426  | 895    | 1101   | 1398   | 100     | 396    | 1802   | 3815   | 2704   | 3203   | 3186   | 3002   |
|      | 3610   | 3404   | 3700   | 2202   | 2700   | 28324   | 5400   |        |        |        |        |        |        |
| L714 | 8142   | 29688  | 5953   | 3644   | 2859   | 411     | 208    | 2568   | 2908   | 3084   | 4299   | 4211   | 4702   |
|      | 6628   | 7349   | 6478   | 2919   | 5343   | 43096   | 7601   |        |        |        |        |        |        |
| L724 | 1038   | 2709   | 300    | 387    | 218    | 0       | 0      | 94     | 397    | 410    | 491    | 334    | 204    |
|      | 296    | 790    | 294    | 808    | 534    | 3968    | 684    |        |        |        |        |        |        |
| L734 | 1954   | 1935   | 135    | 100    | 135    | 0       | 98     | 98     | 371    | 412    | 102    | 422    | 96     |
|      | 242    | 773    | 562    | 399    | 684    | 3063    | 1196   |        |        |        |        |        |        |
| L74  | 11134  | 34332  | 6388   | 4131   | 3212   | 411     | 306    | 2760   | 3676   | 3906   | 4892   | 4967   | 5002   |
|      | 7166   | 8912   | 7334   | 4126   | 5561   | 50127   | 9481   |        |        |        |        |        |        |
| L715 | 112765 | 747477 | 299427 | 164195 | 82080  | 630     | 5357   | 135499 | 140880 | 151355 | 138544 | 138316 | 139552 |
|      | 120606 | 120254 | 114680 | 84553  | 116978 | 1105370 | 301204 |        |        |        |        |        |        |
| L725 | 5520   | 12769  | 958    | 1096   | 980    | 110     | 258    | 753    | 1508   | 1589   | 1271   | 1005   | 2481   |
|      | 264    | 2704   | 2137   | 1720   | 3360   | 18714   | 2719   |        |        |        |        |        |        |
| L735 | 20911  | 12881  | 255    | 366    | 1323   | 0       | 164    | 4940   | 4805   | 3672   | 3189   | 1966   | 1950   |
|      | 3715   | 2500   | 2723   | 2576   | 3536   | 27965   | 7771   |        |        |        |        |        |        |
| L75  | 139196 | 773127 | 300640 | 165657 | 84383  | 740     | 5779   | 141192 | 147193 | 156616 | 143004 | 141287 | 143983 |
|      | 124585 | 125458 | 119540 | 88849  | 123874 | 1152049 | 311694 |        |        |        |        |        |        |
| L811 | 3882   | 11401  | 6835   | 5848   | 2704   | 64      | 16     | 150    | 30     | 343    | 865    | 1144   | 3343   |
|      | 5605   | 6665   | 6864   | 5571   | 138    | 29429   | 1305   |        |        |        |        |        |        |
| L821 | 877    | 1320   | 401    | 601    | 705    | 52      | 0      | 27     | 0      | 131    | 220    | 117    | 194    |
|      | 388    | 497    | 834    | 1454   | 94     | 3679    | 277    |        |        |        |        |        |        |
| L831 | 480    | 658    | 233    | 41     | 308    | 0       | 0      | 167    | 151    | 82     | 199    | 128    | 65     |
|      | 57     | 153    | 258    | 460    | 0      | 1485    | 235    |        |        |        |        |        |        |
| L81  | 5239   | 13379  | 7469   | 6490   | 3717   | 116     | 16     | 344    | 181    | 556    | 1284   | 1389   | 3602   |
|      | 6050   | 7315   | 7956   | 7485   | 232    | 34593   | 1817   |        |        |        |        |        |        |
| L812 | 3485   | 10892  | 4555   | 2063   | 1296   | 90      | 290    | 203    | 193    | 472    | 1092   | 2446   | 2900   |
|      | 4448   | 3788   | 3505   | 2951   | 93     | 19803   | 2578   |        |        |        |        |        |        |
| L822 | 5052   | 11533  | 2582   | 1389   | 3405   | 0       | 364    | 150    | 694    | 738    | 1916   | 2056   | 1722   |
|      | 2219   | 2957   | 3351   | 7621   | 173    | 20775   | 3186   |        |        |        |        |        |        |
| L832 | 9192   | 6328   | 811    | 447    | 746    | 0       | 379    | 310    | 4326   | 903    | 2225   | 1055   | 1420   |
|      | 1426   | 1431   | 1308   | 2501   | 240    | 14763   | 2761   |        |        |        |        |        |        |
| L82  | 17729  | 28753  | 7948   | 3899   | 5447   | 90      | 1033   | 663    | 5213   | 2113   | 5233   | 5557   | 6042   |
|      | 8093   | 9176   | 8164   | 13073  | 506    | 55341   | 8525   |        |        |        |        |        |        |
| L813 | 145    | 523    | 341    | 50     | 120    | 0       | 0      | 168    | 145    | 83     | 172    | 148    | 92     |
|      | 47     | 124    | 116    | 84     | 0      | 1059    | 120    |        |        |        |        |        |        |
| L823 | 3653   | 9668   | 1846   | 306    | 880    | 28      | 318    | 1460   | 2246   | 1499   | 1493   | 1862   | 1662   |
|      | 1726   | 1766   | 1178   | 1502   | 169    | 15268   | 1613   |        |        |        |        |        |        |

APÊNDICE B

B.6 PROGRAMA DO SAS PARA PONTOS SUPLEMENTARES

|      |        |        |         |        |        |         |        |        |        |        |        |        |        |
|------|--------|--------|---------|--------|--------|---------|--------|--------|--------|--------|--------|--------|--------|
| L933 | 13234  | 11805  | 1664    | 790    | 597    | 0       | 30     | 3111   | 2222   | 2492   | 2500   | 2977   | 2969   |
|      | 2631   | 2853   | 3262    | 1682   | 361    | 14250   | 3840   |        |        |        |        |        |        |
| L83  | 17032  | 21996  | 3851    | 1646   | 1597   | 28      | 348    | 4739   | 4613   | 4074   | 4165   | 4987   | 4723   |
|      | 4404   | 4743   | 4556    | 4268   | 530    | 40577   | 5573   |        |        |        |        |        |        |
| L814 | 10963  | 55515  | 34431   | 15838  | 4188   | 689     | 433    | 5573   | 6750   | 9302   | 10227  | 11938  | 15631  |
|      | 20180  | 20487  | 16422   | 7693   | 1988   | 118485  | 8139   |        |        |        |        |        |        |
| L824 | 1426   | 5697   | 1295    | 590    | 606    | 27      | 0      | 341    | 470    | 762    | 962    | 505    | 1105   |
|      | 1298   | 1053   | 1618    | 1314   | 123    | 9031    | 520    |        |        |        |        |        |        |
| L834 | 2954   | 2928   | 483     | 219    | 160    | 0       | 46     | 697    | 592    | 471    | 812    | 529    | 643    |
|      | 626    | 393    | 549     | 1261   | 25     | 5662    | 982    |        |        |        |        |        |        |
| L84  | 17343  | 66950  | 36209   | 16647  | 4954   | 716     | 479    | 5611   | 7812   | 10535  | 12001  | 12972  | 17379  |
|      | 12104  | 21933  | 18589   | 10268  | 2136   | 133178  | 9641   |        |        |        |        |        |        |
| L815 | 88619  | 641006 | 779697  | 383738 | 92467  | 9540    | 3393   | 202330 | 212046 | 238489 | 230004 | 215452 | 217239 |
|      | 198121 | 177532 | 157598  | 122810 | 20053  | 1762342 | 232725 |        |        |        |        |        |        |
| L825 | 5454   | 23195  | 6857    | 3501   | 2203   | 191     | 18     | 3419   | 2776   | 3877   | 4267   | 4497   | 4565   |
|      | 5208   | 4658   | 4759    | 3995   | 462    | 38918   | 2483   |        |        |        |        |        |        |
| L835 | 19907  | 15085  | 802     | 757    | 537    | 0       | 0      | 4854   | 4071   | 4071   | 3859   | 3227   | 3076   |
|      | 3280   | 3682   | 3295    | 3418   | 255    | 32404   | 4684   |        |        |        |        |        |        |
| L85  | 113980 | 679296 | 787356  | 387996 | 95207  | 9731    | 3411   | 209603 | 218893 | 246437 | 238130 | 223176 | 224880 |
|      | 206609 | 185872 | 165652  | 130123 | 20770  | 1833664 | 239892 |        |        |        |        |        |        |
| L911 | 6996   | 33812  | 75958   | 38539  | 9382   | 271     | 1687   | 12854  | 13957  | 16244  | 16309  | 17893  | 17417  |
|      | 16572  | 15033  | 16389   | 14241  | 6362   | 153093  | 11866  |        |        |        |        |        |        |
| L921 | 909    | 4950   | 2907    | 1962   | 1454   | 0       | 178    | 666    | 743    | 1115   | 1131   | 981    | 1102   |
|      | 1179   | 1424   | 1374    | 1831   | 456    | 11713   | 469    |        |        |        |        |        |        |
| L931 | 2439   | 3377   | 359     | 193    | 130    | 0       | 52     | 667    | 582    | 597    | 553    | 656    | 530    |
|      | 675    | 519    | 791     | 732    | 247    | 6174    | 425    |        |        |        |        |        |        |
| L91  | 10344  | 42139  | 79224   | 40694  | 11066  | 271     | 1917   | 14187  | 15282  | 17956  | 17993  | 19530  | 19049  |
|      | 18426  | 16976  | 18554   | 16904  | 7065   | 170980  | 12760  |        |        |        |        |        |        |
| L912 | 2012   | 6749   | 8164    | 4542   | 2578   | 10      | 1362   | 130    | 703    | 656    | 1195   | 2285   | 2091   |
|      | 3375   | 3789   | 4371    | 3743   | 356    | 21757   | 2299   |        |        |        |        |        |        |
| L922 | 5775   | 20839  | 8026    | 6986   | 11251  | 0       | 1613   | 646    | 2394   | 1704   | 1927   | 3938   | 2847   |
|      | 4504   | 5363   | 7513    | 19484  | 944    | 49042   | 3835   |        |        |        |        |        |        |
| L932 | 9580   | 11871  | 1422    | 929    | 1280   | 0       | 436    | 1597   | 5608   | 1274   | 1588   | 3436   | 1551   |
|      | 1961   | 2032   | 2022    | 3323   | 254    | 22287   | 2795   |        |        |        |        |        |        |
| L92  | 17367  | 39459  | 17612   | 12457  | 15109  | 10      | 3411   | 2373   | 8705   | 3634   | 4710   | 9659   | 6489   |
|      | 9840   | 11184  | 13906   | 26550  | 1554   | 93086   | 8929   |        |        |        |        |        |        |
| L913 | 419    | 1903   | 2751    | 1509   | 498    | 21      | 254    | 199    | 388    | 653    | 567    | 643    | 600    |
|      | 928    | 1154   | 955     | 486    | 174    | 6729    | 272    |        |        |        |        |        |        |
| L923 | 888    | 3724   | 1500    | 892    | 998    | 0       | 154    | 690    | 654    | 471    | 715    | 897    | 558    |
|      | 772    | 820    | 911     | 1162   | 198    | 7614    | 388    |        |        |        |        |        |        |
| L933 | 2023   | 3303   | 434     | 310    | 210    | 0       | 28     | 563    | 493    | 607    | 592    | 556    | 502    |
|      | 548    | 752    | 733     | 728    | 177    | 5861    | 420    |        |        |        |        |        |        |
| L93  | 3330   | 8830   | 4685    | 2711   | 1706   | 21      | 436    | 1452   | 1535   | 1731   | 1874   | 2096   | 1660   |
|      | 2248   | 2726   | 2599    | 2376   | 549    | 20204   | 1080   |        |        |        |        |        |        |
| L914 | 3592   | 17664  | 44286   | 21392  | 5080   | 162     | 848    | 6145   | 7236   | 9303   | 8910   | 10524  | 9316   |
|      | 9414   | 10291  | 9325    | 7083   | 3780   | 86331   | 5844   |        |        |        |        |        |        |
| L924 | 585    | 2438   | 1572    | 786    | 889    | 0       | 62     | 431    | 543    | 418    | 433    | 611    | 634    |
|      | 665    | 600    | 845     | 785    | 243    | 5994    | 276    |        |        |        |        |        |        |
| L934 | 1455   | 2816   | 124     | 158    | 130    | 0       | 3      | 311    | 323    | 319    | 301    | 334    | 250    |
|      | 430    | 342    | 440     | 451    | 179    | 3424    | 259    |        |        |        |        |        |        |
| L94  | 5632   | 21918  | 45982   | 22336  | 6099   | 162     | 913    | 6887   | 8102   | 10040  | 9644   | 11469  | 10200  |
|      | 10509  | 11233  | 10610   | 8319   | 4202   | 95749   | 6379   |        |        |        |        |        |        |
| L915 | 90571  | 469510 | 1017229 | 461993 | 103592 | 3264    | 17414  | 178307 | 192848 | 228002 | 229907 | 234517 | 218549 |
|      | 221890 | 206399 | 187612  | 143175 | 87540  | 1983937 | 162222 |        |        |        |        |        |        |
| L925 | 11806  | 61900  | 32513   | 18777  | 13655  | 59      | 1023   | 9654   | 9709   | 11875  | 12992  | 14584  | 13587  |
|      | 14192  | 15188  | 15603   | 14624  | 5679   | 132451  | 6259   |        |        |        |        |        |        |
| L935 | 30293  | 36166  | 3700    | 2179   | 2125   | 0       | 156    | 7520   | 7336   | 7349   | 6795   | 7035   | 6599   |
|      | 6912   | 7504   | 7573    | 7461   | 2223   | 68178   | 6285   |        |        |        |        |        |        |
| L95  | 132670 | 567576 | 1053442 | 482949 | 119372 | 3323    | 18593  | 195481 | 209893 | 247226 | 249694 | 256136 | 238735 |
|      | 242994 | 229091 | 210788  | 165260 | 95442  | 2184566 | 174766 |        |        |        |        |        |        |

```

;
RUN;
proc transpose data=HLS2T out=HLs2TT;
  VAR  i0 i1 i2 i3 i4 in r0 r1 r2 r3 r4 r5 r6 r7 r8 r9 r10 rn zu zr;
  ID   LOCAL;
RUN;

proc corresp data=HLS2TT out=SHLS2TT ALL profile=column;
  VAR L711 L721 L731 L71 L712 L722 L732 L72 L713 L723 L733 L73 L714 L724 L734 L74 L715 L725 L735
      L75
      L811 L821 L831 L81 L812 L822 L832 L82 L813 L823 L833 L83 L814 L824 L834 L84 L815 L825 L835
      L85
      L911 L921 L931 L91 L912 L922 L932 L92 L913 L923 L933 L93 L914 L924 L934 L94 L915 L925 L935
      L95;

```

```
supplementary
  L711 L712 L713 L714 L715 L721 L722 L723 L724 L725 L731 L732 L733 L734 L735
  L811 L812 L813 L814 L815 L821 L822 L823 L824 L825 L831 L832 L833 L834 L835
  L911 L912 L913 L914 L915 L921 L922 L923 L924 L925 L931 L932 L933 L934 L935;

  ID _NAME_;
run;

DATA GHLS2T;
  SET GHLS2TT;
  X=DIM1;
  Y=DIM2;
  XSYS='2';
  YSYS='2';
  TEXT= _NAME_;
  SIZE=1.5;
  LABEL X='DIMENSAO 1'
        Y='DIMENSAO 2';
  KEEP X Y TEXT XSYS YSYS SIZE;
run;

PROC PRINT DATA=GHLS2TT;
run;

FILENAME FIG 'C:\USERS\ADMUR\pasta\GHLS2TS.GRF';
GOPTIONS DEVICE=CGMWPWA GSFNAME=FIG GSFMODE=REPLACE;

PROC GPLOT DATA=GHLS2TT;
  SYMBOL VALUE=NONE;
  AXIS1 ORDER=-1 TO 8 BY 1;
  PLOT Y*X=1 / ANNOTATE=GHLS2TT
           VAXIS=AXIS1 HAXIS=AXIS1
           VREF=0 HREF=0
           FRAME;
run;

GOPTIONS DEVICE=WIN;
QUIT;

FILENAME FIG 'C:\USERS\ADMUR\pasta\GHLS2TZ.GRF';
GOPTIONS DEVICE=CGMWPWA GSFNAME=FIG GSFMODE=REPLACE;

PROC GPLOT DATA=GHLS2TT;
  SYMBOL VALUE=NONE;
  AXIS1 ORDER=-0.4 TO 1.2 BY 0.4;
  PLOT Y*X=1 / ANNOTATE=GHLS2TT
           VAXIS=AXIS1 HAXIS=AXIS1
           VREF=0 HREF=0
           FRAME;
run;

GOPTIONS DEVICE=WIN;
QUIT;
```

**B.7 SAÍDA DO PROGRAMA/SAS PARA OS PONTOS SUPLEMENTARES**

The SAS System  
The Correspondence Analysis Procedure  
Contingency Table

|     | L71   | L72    | L73    | L74    | L75     | L81    | L82    | L83    |
|-----|-------|--------|--------|--------|---------|--------|--------|--------|
| I0  | 1376  | 17412  | 12804  | 11134  | 139196  | 5239   | 17729  | 17032  |
| I1  | 2937  | 17308  | 17426  | 34332  | 773127  | 13379  | 29753  | 21996  |
| I2  | 947   | 1887   | 895    | 6388   | 300640  | 7469   | 7948   | 3851   |
| I3  | 924   | 2314   | 1101   | 4131   | 165657  | 6490   | 3899   | 1646   |
| I4  | 3731  | 3588   | 1398   | 3212   | 84383   | 3717   | 5447   | 1597   |
| IN  | 0     | 0      | 100    | 411    | 740     | 116    | 90     | 28     |
| R0  | 1362  | 2599   | 396    | 306    | 5779    | 16     | 1033   | 348    |
| R1  | 117   | 102    | 1802   | 2760   | 141192  | 344    | 663    | 4739   |
| R2  | 0     | 1510   | 1815   | 3676   | 147193  | 181    | 5213   | 4613   |
| R3  | 520   | 1105   | 2704   | 3906   | 156616  | 556    | 2113   | 4074   |
| R4  | 235   | 1193   | 3203   | 4892   | 143004  | 1284   | 5233   | 4165   |
| R5  | 217   | 3107   | 3186   | 4967   | 141287  | 1389   | 5557   | 4987   |
| R6  | 943   | 2711   | 3002   | 5002   | 143983  | 3602   | 6042   | 4723   |
| R7  | 1119  | 7398   | 3610   | 7166   | 124585  | 6050   | 8093   | 4404   |
| R8  | 1075  | 7491   | 3404   | 8912   | 125458  | 7315   | 8176   | 4743   |
| R9  | 1738  | 6314   | 3700   | 7334   | 119540  | 7956   | 8164   | 4556   |
| R10 | 1945  | 7479   | 2202   | 4126   | 88849   | 7485   | 13073  | 4268   |
| RN  | 644   | 1500   | 2700   | 6561   | 123874  | 232    | 506    | 530    |
| ZU  | 8854  | 31047  | 28324  | 50127  | 1152049 | 34593  | 53341  | 40577  |
| ZR  | 1061  | 11462  | 5400   | 9481   | 311694  | 1817   | 9525   | 5573   |
| Sum | 29745 | 127527 | 101172 | 178824 | 4388846 | 109230 | 191598 | 138450 |

|     | L84    | L85     | L91    | L92    | L93   | L94    | L95     | Sum      |
|-----|--------|---------|--------|--------|-------|--------|---------|----------|
| I0  | 17343  | 113980  | 10344  | 17367  | 3330  | 5632   | 132670  | 522588   |
| I1  | 66950  | 679286  | 42139  | 39459  | 8830  | 21218  | 567576  | 2335416  |
| I2  | 36209  | 787356  | 73224  | 17612  | 4685  | 45982  | 1053442 | 2354535  |
| I3  | 16647  | 387996  | 40694  | 12457  | 2711  | 22336  | 482949  | 1151952  |
| I4  | 4954   | 35207   | 11066  | 15109  | 1706  | 6099   | 119372  | 360586   |
| IN  | 716    | 9731    | 271    | 10     | 21    | 162    | 3323    | 15719    |
| R0  | 479    | 3411    | 1917   | 3411   | 436   | 913    | 18593   | 40999    |
| R1  | 6611   | 209603  | 14187  | 2373   | 1452  | 6887   | 195481  | 588313   |
| R2  | 7812   | 218893  | 15282  | 8705   | 1535  | 8102   | 209893  | 636423   |
| R3  | 10535  | 246437  | 17956  | 3634   | 1731  | 10040  | 247226  | 709153   |
| R4  | 12001  | 238130  | 17993  | 4710   | 1874  | 9644   | 249694  | 697255   |
| R5  | 12972  | 223176  | 19530  | 9659   | 2096  | 11469  | 256136  | 699735   |
| R6  | 17379  | 224880  | 19049  | 6489   | 1660  | 10200  | 238735  | 688400   |
| R7  | 22104  | 206609  | 18426  | 9840   | 2248  | 10509  | 242994  | 675155   |
| R8  | 21933  | 185872  | 16976  | 11184  | 2726  | 11233  | 229091  | 645589   |
| R9  | 18589  | 165652  | 18554  | 13906  | 2599  | 10610  | 210788  | 600000   |
| R10 | 10268  | 130123  | 16804  | 26550  | 2376  | 8319   | 165260  | 489127   |
| RN  | 2136   | 20770   | 7065   | 1554   | 549   | 4202   | 95442   | 268265   |
| ZU  | 133178 | 1833664 | 170980 | 93086  | 20204 | 35749  | 2184566 | 5932339  |
| ZR  | 9641   | 239892  | 12760  | 8929   | 1080  | 6379   | 174766  | 308460   |
| Sum | 428457 | 6220668 | 531217 | 306044 | 63849 | 356385 | 7077997 | 20220009 |

Supplementary Columns

|     | L711 | L721 | L731 | L712  | L722  | L732 | L713 | L723 | L733  |
|-----|------|------|------|-------|-------|------|------|------|-------|
| I0  | 752  | 503  | 121  | 5200  | 5398  | 6814 | 401  | 2802 | 9601  |
| I1  | 1822 | 697  | 418  | 6194  | 7208  | 3906 | 1205 | 6021 | 10200 |
| I2  | 947  | 0    | 0    | 1289  | 498   | 100  | 100  | 598  | 197   |
| I3  | 924  | 0    | 0    | 1105  | 1108  | 101  | 0    | 801  | 300   |
| I4  | 3413 | 218  | 100  | 1489  | 1800  | 299  | 297  | 700  | 401   |
| IN  | 0    | 0    | 0    | 0     | 0     | 0    | 0    | 0    | 100   |
| R0  | 1362 | 0    | 0    | 392   | 198   | 2009 | 0    | 99   | 297   |
| R1  | 117  | 0    | 0    | 0     | 0     | 102  | 403  | 503  | 896   |
| R2  | 0    | 0    | 0    | 308   | 303   | 899  | 405  | 1213 | 2197  |
| R3  | 420  | 100  | 0    | 305   | 598   | 202  | 99   | 1204 | 1401  |
| R4  | 0    | 235  | 0    | 398   | 199   | 596  | 199  | 1501 | 1503  |
| R5  | 0    | 0    | 217  | 495   | 1399  | 1213 | 297  | 598  | 2291  |
| R6  | 943  | 0    | 0    | 707   | 505   | 1499 | 0    | 1003 | 1999  |
| R7  | 626  | 392  | 101  | 3703  | 2303  | 1392 | 302  | 1202 | 2106  |
| R8  | 954  | 0    | 121  | 3288  | 3099  | 1104 | 100  | 799  | 2505  |
| R9  | 1638 | 0    | 100  | 2998  | 2311  | 1005 | 198  | 1404 | 2098  |
| R10 | 1271 | 574  | 100  | 2485  | 4196  | 798  | 0    | 500  | 1702  |
| RN  | 527  | 117  | 0    | 198   | 901   | 401  | 0    | 896  | 1804  |
| ZU  | 7100 | 1115 | 639  | 11759 | 12185 | 7103 | 1700 | 9409 | 17215 |
| ZR  | 758  | 303  | 0    | 3518  | 3827  | 4117 | 303  | 1513 | 3584  |

|    | L714  | L724 | L734 | L715   | L725  | L735  | L811  | L821 | L831 |
|----|-------|------|------|--------|-------|-------|-------|------|------|
| I0 | 8142  | 1038 | 1954 | 112765 | 5520  | 20911 | 3882  | 877  | 480  |
| I1 | 29688 | 2709 | 1935 | 747477 | 12769 | 12881 | 11401 | 1320 | 658  |
| I2 | 5953  | 300  | 135  | 299427 | 958   | 255   | 6835  | 401  | 233  |

APÊNDICE B

B.7 SAÍDA DO PROGRAMA/SAS DOS PONTOS SUPLEMENTARES

|     |         |       |       |         |       |       |         |        |       |
|-----|---------|-------|-------|---------|-------|-------|---------|--------|-------|
| I3  | 3644    | 387   | 100   | 164195  | 1096  | 366   | 5848    | 601    | 41    |
| I4  | 2859    | 218   | 135   | 82080   | 980   | 1323  | 2704    | 705    | 308   |
| IN  | 411     | 0     | 0     | 630     | 110   | 0     | 64      | 52     | 0     |
| RO  | 208     | 0     | 98    | 5357    | 258   | 164   | 16      | 0      | 0     |
| R1  | 2568    | 94    | 98    | 135499  | 753   | 4940  | 150     | 27     | 167   |
| R2  | 2908    | 397   | 371   | 140880  | 1508  | 4805  | 30      | 0      | 151   |
| R3  | 3084    | 410   | 412   | 151355  | 1589  | 3672  | 343     | 131    | 82    |
| R4  | 4299    | 491   | 102   | 138544  | 1271  | 3189  | 865     | 220    | 199   |
| R5  | 4211    | 334   | 422   | 138316  | 1005  | 1966  | 1144    | 117    | 128   |
| R6  | 4702    | 204   | 96    | 139552  | 2481  | 1950  | 3343    | 194    | 65    |
| R7  | 6628    | 296   | 242   | 120606  | 264   | 3715  | 5605    | 388    | 57    |
| R8  | 7349    | 790   | 773   | 120254  | 2704  | 2500  | 6665    | 497    | 153   |
| R9  | 6478    | 294   | 562   | 114680  | 2137  | 2723  | 6864    | 834    | 258   |
| R10 | 2919    | 908   | 599   | 84553   | 1720  | 2576  | 5571    | 1454   | 460   |
| RN  | 5343    | 534   | 684   | 116978  | 3360  | 3536  | 138     | 94     | 0     |
| ZU  | 43096   | 3968  | 3063  | 1105370 | 18714 | 27965 | 19429   | 3679   | 1485  |
| ZR  | 7601    | 684   | 1196  | 391204  | 2719  | 7771  | 1305    | 277    | 235   |
|     | L812    | L822  | L832  | L813    | L823  | L833  | L814    | L824   | L834  |
| I0  | 3485    | 5052  | 9192  | 145     | 3653  | 13234 | 12963   | 1426   | 2954  |
| I1  | 10892   | 11533 | 6328  | 523     | 9668  | 11805 | 98515   | 5607   | 2828  |
| I2  | 4555    | 2582  | 911   | 341     | 1846  | 1664  | 34431   | 1295   | 483   |
| I3  | 2063    | 1389  | 447   | 50      | 806   | 790   | 15838   | 590    | 219   |
| I4  | 1296    | 3405  | 746   | 120     | 880   | 597   | 4188    | 606    | 160   |
| IN  | 90      | 0     | 0     | 0       | 28    | 0     | 689     | 27     | 0     |
| RO  | 290     | 364   | 379   | 0       | 318   | 30    | 433     | 0      | 46    |
| R1  | 203     | 150   | 310   | 168     | 1460  | 3111  | 5573    | 341    | 697   |
| R2  | 193     | 694   | 4326  | 145     | 2246  | 2222  | 6750    | 470    | 592   |
| R3  | 472     | 738   | 503   | 83      | 1499  | 2492  | 3302    | 762    | 471   |
| R4  | 1092    | 1316  | 2225  | 172     | 1493  | 2500  | 10227   | 362    | 812   |
| R5  | 2446    | 2056  | 1055  | 148     | 1862  | 2977  | 11938   | 505    | 529   |
| R6  | 2900    | 1722  | 1420  | 92      | 1662  | 2969  | 15631   | 1105   | 643   |
| R7  | 4448    | 2219  | 1426  | 47      | 1726  | 2631  | 20180   | 1298   | 626   |
| R8  | 3788    | 2957  | 1431  | 124     | 1766  | 2853  | 20487   | 1053   | 393   |
| R9  | 3505    | 2351  | 1308  | 116     | 1178  | 3262  | 16422   | 1618   | 549   |
| R10 | 2951    | 7621  | 2501  | 84      | 1502  | 2682  | 7693    | 1314   | 1261  |
| RN  | 93      | 173   | 240   | 0       | 169   | 361   | 1988    | 123    | 25    |
| ZU  | 19803   | 20775 | 14763 | 1059    | 15268 | 24250 | 118485  | 9031   | 5662  |
| ZR  | 2578    | 3186  | 2761  | 120     | 1613  | 3840  | 8139    | 520    | 982   |
|     | L815    | L825  | L835  | L911    | L921  | L931  | L912    | L922   | L932  |
| I0  | 88619   | 6454  | 19907 | 6996    | 909   | 2439  | 2012    | 5775   | 9580  |
| I1  | 641006  | 23195 | 15085 | 33812   | 4950  | 3377  | 6749    | 20839  | 11871 |
| I2  | 779697  | 6857  | 802   | 75958   | 2907  | 359   | 8164    | 8026   | 1422  |
| I3  | 383738  | 3501  | 757   | 38539   | 1962  | 193   | 4542    | 6986   | 929   |
| I4  | 92467   | 2203  | 537   | 9382    | 1454  | 230   | 2578    | 11251  | 1280  |
| IN  | 9540    | 191   | 0     | 271     | 0     | 0     | 10      | 0      | 0     |
| RO  | 3393    | 18    | 0     | 1687    | 178   | 52    | 1362    | 1613   | 436   |
| R1  | 202330  | 2419  | 4854  | 12854   | 666   | 667   | 130     | 646    | 1597  |
| R2  | 212046  | 2776  | 4071  | 13957   | 743   | 582   | 703     | 2394   | 5608  |
| R3  | 238489  | 3877  | 4071  | 16244   | 1115  | 597   | 656     | 1704   | 1274  |
| R4  | 230004  | 4267  | 3659  | 16309   | 1131  | 553   | 1195    | 1927   | 1588  |
| R5  | 215452  | 4497  | 3227  | 17893   | 381   | 656   | 2285    | 3938   | 3436  |
| R6  | 217239  | 4565  | 3076  | 17417   | 1102  | 530   | 2091    | 2847   | 1551  |
| R7  | 198121  | 5208  | 3280  | 16572   | 1179  | 675   | 3375    | 4504   | 1961  |
| R8  | 177532  | 4658  | 3682  | 15033   | 1424  | 519   | 3789    | 5363   | 2032  |
| R9  | 157598  | 4759  | 3295  | 16389   | 1374  | 791   | 4371    | 7513   | 2022  |
| R10 | 122810  | 3895  | 3418  | 14241   | 1831  | 732   | 3743    | 19484  | 3323  |
| RN  | 20053   | 462   | 255   | 6362    | 456   | 247   | 356     | 944    | 254   |
| ZU  | 1762342 | 38918 | 32404 | 153093  | 11713 | 6174  | 21757   | 49042  | 22287 |
| ZR  | 232725  | 2483  | 4684  | 11866   | 469   | 425   | 2299    | 3835   | 2795  |
|     | L913    | L923  | L933  | L914    | L924  | L934  | L915    | L925   | L935  |
| I0  | 419     | 888   | 2023  | 3592    | 585   | 1455  | 90571   | 11806  | 30293 |
| I1  | 1803    | 3724  | 3303  | 17664   | 2438  | 1816  | 469510  | 61900  | 36166 |
| I2  | 2751    | 1500  | 434   | 44286   | 1572  | 124   | 1017229 | 32513  | 3700  |
| I3  | 1509    | 892   | 310   | 21392   | 786   | 158   | 461993  | 18777  | 2179  |
| I4  | 498     | 998   | 210   | 5080    | 889   | 130   | 103592  | 13655  | 2125  |
| IN  | 21      | 0     | 0     | 162     | 0     | 0     | 3264    | 59     | 0     |
| RO  | 254     | 154   | 28    | 848     | 62    | 3     | 17414   | 1023   | 156   |
| R1  | 199     | 690   | 563   | 6145    | 431   | 311   | 178307  | 9654   | 7520  |
| R2  | 388     | 654   | 493   | 7236    | 543   | 323   | 192848  | 9709   | 7336  |
| R3  | 653     | 471   | 607   | 9303    | 418   | 319   | 228002  | 11875  | 7349  |
| R4  | 567     | 715   | 592   | 8910    | 433   | 301   | 229907  | 12992  | 6795  |
| R5  | 643     | 897   | 556   | 10524   | 611   | 334   | 234517  | 14584  | 7035  |
| R6  | 600     | 558   | 502   | 9316    | 634   | 250   | 218549  | 13587  | 6599  |
| R7  | 928     | 772   | 548   | 9414    | 665   | 430   | 221890  | 14192  | 6912  |
| R8  | 1154    | 820   | 752   | 10291   | 600   | 342   | 206399  | 15188  | 7504  |
| R9  | 955     | 911   | 733   | 9325    | 845   | 440   | 187612  | 15603  | 7573  |
| R10 | 486     | 1162  | 728   | 7083    | 785   | 451   | 143175  | 14624  | 7461  |
| RN  | 174     | 198   | 177   | 3780    | 243   | 179   | 87540   | 5679   | 2223  |
| ZU  | 6729    | 7614  | 5861  | 86331   | 5994  | 3424  | 1983937 | 132451 | 68178 |
| ZR  | 272     | 388   | 420   | 5844    | 276   | 259   | 162222  | 6259   | 6285  |

Row Profiles

|     | L71      | L72      | L73      | L74      | L75      | L81      | L82      | L83      |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|
| I0  | 0.002633 | 0.033319 | 0.024501 | 0.021306 | 0.266359 | 0.010025 | 0.033925 | 0.032592 |
| I1  | 0.001258 | 0.007411 | 0.007462 | 0.014701 | 0.331045 | 0.005729 | 0.012312 | 0.009418 |
| I2  | 0.000402 | 0.000801 | 0.000380 | 0.002713 | 0.127686 | 0.003172 | 0.003376 | 0.001636 |
| I3  | 0.000802 | 0.002009 | 0.000956 | 0.003586 | 0.143805 | 0.005634 | 0.003385 | 0.001429 |
| I4  | 0.010347 | 0.009950 | 0.003877 | 0.008908 | 0.234016 | 0.010308 | 0.015106 | 0.004429 |
| IN  | 0.000000 | 0.000000 | 0.006362 | 0.026147 | 0.047077 | 0.007380 | 0.005726 | 0.001781 |
| RO  | 0.033220 | 0.063392 | 0.009659 | 0.007464 | 0.140955 | 0.000390 | 0.025196 | 0.008488 |
| R1  | 0.000199 | 0.000173 | 0.003063 | 0.004691 | 0.239995 | 0.000585 | 0.001127 | 0.008055 |
| R2  | 0.000000 | 0.002373 | 0.005994 | 0.005776 | 0.231292 | 0.000284 | 0.008191 | 0.007248 |
| R3  | 0.000733 | 0.001558 | 0.003813 | 0.005508 | 0.220849 | 0.000784 | 0.002980 | 0.005745 |
| R4  | 0.000337 | 0.001711 | 0.004594 | 0.007016 | 0.205096 | 0.001842 | 0.007505 | 0.005973 |
| R5  | 0.000310 | 0.004440 | 0.004553 | 0.007098 | 0.201915 | 0.001985 | 0.007942 | 0.007127 |
| R6  | 0.001370 | 0.003938 | 0.004361 | 0.007266 | 0.209156 | 0.005232 | 0.008777 | 0.006861 |
| R7  | 0.001657 | 0.010957 | 0.005347 | 0.010614 | 0.184528 | 0.008961 | 0.011987 | 0.006523 |
| R8  | 0.001665 | 0.011603 | 0.005273 | 0.013804 | 0.194331 | 0.011331 | 0.012664 | 0.007347 |
| R9  | 0.002897 | 0.010523 | 0.006167 | 0.012223 | 0.199233 | 0.013260 | 0.013607 | 0.007593 |
| R10 | 0.003976 | 0.015291 | 0.004502 | 0.008435 | 0.181648 | 0.015303 | 0.026727 | 0.008726 |
| RN  | 0.002401 | 0.005591 | 0.010065 | 0.024457 | 0.461760 | 0.000865 | 0.001886 | 0.001976 |
| ZU  | 0.001492 | 0.005234 | 0.004775 | 0.008450 | 0.194198 | 0.005831 | 0.009329 | 0.006840 |
| ZR  | 0.001312 | 0.014178 | 0.006679 | 0.011727 | 0.385540 | 0.002247 | 0.010545 | 0.006893 |

|     | L84      | L85      | L91      | L92      | L93      | L94      | L95      |
|-----|----------|----------|----------|----------|----------|----------|----------|
| I0  | 0.033187 | 0.218107 | 0.019794 | 0.033233 | 0.006372 | 0.010777 | 0.253871 |
| I1  | 0.028667 | 0.290863 | 0.018043 | 0.016896 | 0.003781 | 0.009385 | 0.243030 |
| I2  | 0.015378 | 0.334400 | 0.033647 | 0.007480 | 0.001990 | 0.019529 | 0.447410 |
| I3  | 0.014451 | 0.336816 | 0.035326 | 0.010814 | 0.002353 | 0.019390 | 0.419244 |
| I4  | 0.013739 | 0.264034 | 0.030689 | 0.041901 | 0.004731 | 0.016914 | 0.331050 |
| IN  | 0.045550 | 0.619060 | 0.017240 | 0.000636 | 0.001336 | 0.010306 | 0.211400 |
| RO  | 0.011683 | 0.083197 | 0.046757 | 0.083197 | 0.010634 | 0.022269 | 0.453499 |
| R1  | 0.011237 | 0.356278 | 0.024115 | 0.004034 | 0.002468 | 0.011706 | 0.332274 |
| R2  | 0.012275 | 0.343943 | 0.024012 | 0.013678 | 0.002412 | 0.012731 | 0.329801 |
| R3  | 0.014856 | 0.347509 | 0.025320 | 0.005124 | 0.002441 | 0.014158 | 0.348622 |
| R4  | 0.017212 | 0.341525 | 0.025805 | 0.006755 | 0.002688 | 0.013831 | 0.358110 |
| R5  | 0.018538 | 0.318944 | 0.027911 | 0.013804 | 0.002995 | 0.016390 | 0.366047 |
| R6  | 0.025245 | 0.326671 | 0.027671 | 0.009426 | 0.002411 | 0.014817 | 0.346797 |
| R7  | 0.032739 | 0.306017 | 0.027292 | 0.014574 | 0.003330 | 0.015565 | 0.359908 |
| R8  | 0.033974 | 0.287911 | 0.026295 | 0.017324 | 0.004223 | 0.017400 | 0.354856 |
| R9  | 0.030982 | 0.276087 | 0.030923 | 0.023177 | 0.004332 | 0.017683 | 0.351313 |
| R10 | 0.020993 | 0.266031 | 0.034355 | 0.054280 | 0.004858 | 0.017008 | 0.337867 |
| RN  | 0.007962 | 0.077423 | 0.026336 | 0.005793 | 0.002046 | 0.015664 | 0.355775 |
| ZU  | 0.022449 | 0.309096 | 0.028822 | 0.015691 | 0.003406 | 0.016140 | 0.368247 |
| ZR  | 0.011925 | 0.296727 | 0.015783 | 0.011044 | 0.001336 | 0.007890 | 0.216171 |

Column Profiles

|     | L71      | L72      | L73      | L74      | L75      | L81      | L82      | L83      |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|
| I0  | 0.046260 | 0.136536 | 0.126557 | 0.062262 | 0.031716 | 0.047963 | 0.092532 | 0.123019 |
| I1  | 0.098739 | 0.135720 | 0.172241 | 0.191988 | 0.176157 | 0.122485 | 0.150069 | 0.158873 |
| I2  | 0.031837 | 0.014797 | 0.008846 | 0.035722 | 0.068501 | 0.068379 | 0.041483 | 0.027815 |
| I3  | 0.031064 | 0.018145 | 0.010882 | 0.023101 | 0.037745 | 0.059416 | 0.020350 | 0.011889 |
| I4  | 0.125433 | 0.028135 | 0.013818 | 0.017962 | 0.019227 | 0.034029 | 0.028429 | 0.011535 |
| IN  | 0.000000 | 0.000000 | 0.000988 | 0.002298 | 0.000169 | 0.001062 | 0.000470 | 0.000202 |
| RO  | 0.045789 | 0.020380 | 0.003914 | 0.001711 | 0.001317 | 0.000146 | 0.005391 | 0.002514 |
| R1  | 0.003933 | 0.000800 | 0.017811 | 0.015434 | 0.032171 | 0.003149 | 0.003460 | 0.034229 |
| R2  | 0.000000 | 0.011841 | 0.037708 | 0.020557 | 0.033538 | 0.001657 | 0.027208 | 0.033319 |
| R3  | 0.017482 | 0.008665 | 0.026727 | 0.021843 | 0.035685 | 0.005090 | 0.011028 | 0.029426 |
| R4  | 0.007900 | 0.009355 | 0.031659 | 0.027357 | 0.032584 | 0.011755 | 0.027312 | 0.030083 |
| R5  | 0.007295 | 0.024363 | 0.031491 | 0.027776 | 0.032192 | 0.012716 | 0.029003 | 0.036020 |
| R6  | 0.031703 | 0.021258 | 0.029672 | 0.027972 | 0.032807 | 0.032976 | 0.031535 | 0.034113 |
| R7  | 0.037620 | 0.058011 | 0.035682 | 0.040073 | 0.028387 | 0.055388 | 0.042239 | 0.031809 |
| R8  | 0.036141 | 0.058741 | 0.033646 | 0.049837 | 0.028586 | 0.066969 | 0.042673 | 0.034258 |
| R9  | 0.058430 | 0.049511 | 0.036571 | 0.041012 | 0.027237 | 0.072837 | 0.042610 | 0.032907 |
| R10 | 0.065389 | 0.058646 | 0.021765 | 0.023073 | 0.020244 | 0.068525 | 0.068231 | 0.030827 |
| RN  | 0.021651 | 0.011762 | 0.026687 | 0.036690 | 0.028225 | 0.002124 | 0.002641 | 0.003828 |
| ZU  | 0.297663 | 0.243454 | 0.279959 | 0.280315 | 0.262495 | 0.316699 | 0.288839 | 0.293081 |
| ZR  | 0.035670 | 0.089879 | 0.053374 | 0.053019 | 0.071020 | 0.016635 | 0.044494 | 0.040253 |

|    | L84      | L85      | L91      | L92      | L93      | L94      | L95      |
|----|----------|----------|----------|----------|----------|----------|----------|
| I0 | 0.040478 | 0.018323 | 0.018766 | 0.056747 | 0.052154 | 0.018382 | 0.018744 |
| I1 | 0.156258 | 0.109198 | 0.076447 | 0.128932 | 0.138295 | 0.071537 | 0.080189 |
| I2 | 0.084510 | 0.126571 | 0.143726 | 0.057547 | 0.073376 | 0.150079 | 0.148833 |
| I3 | 0.038853 | 0.062372 | 0.073826 | 0.040703 | 0.042460 | 0.072902 | 0.068232 |
| I4 | 0.011562 | 0.015305 | 0.020076 | 0.049369 | 0.026719 | 0.019906 | 0.016865 |
| IN | 0.001671 | 0.001564 | 0.000492 | 0.000033 | 0.000329 | 0.000529 | 0.000469 |
| RO | 0.001118 | 0.000548 | 0.003478 | 0.011145 | 0.006829 | 0.002980 | 0.002627 |
| R1 | 0.015430 | 0.033695 | 0.025738 | 0.007754 | 0.022741 | 0.022478 | 0.027619 |
| R2 | 0.018233 | 0.035188 | 0.027724 | 0.028444 | 0.024041 | 0.026444 | 0.029654 |
| R3 | 0.024588 | 0.039616 | 0.032575 | 0.011874 | 0.027111 | 0.032769 | 0.034929 |

|     |          |          |          |          |          |          |          |
|-----|----------|----------|----------|----------|----------|----------|----------|
| R4  | 0.028010 | 0.038280 | 0.032642 | 0.015390 | 0.029350 | 0.031477 | 0.035277 |
| R5  | 0.030276 | 0.025877 | 0.035431 | 0.031561 | 0.032827 | 0.037433 | 0.036188 |
| R6  | 0.040562 | 0.036150 | 0.034558 | 0.021203 | 0.025999 | 0.033291 | 0.033729 |
| R7  | 0.051590 | 0.033213 | 0.033428 | 0.032152 | 0.035208 | 0.034300 | 0.034331 |
| R8  | 0.051191 | 0.029880 | 0.030797 | 0.036544 | 0.042694 | 0.036663 | 0.032367 |
| R9  | 0.043386 | 0.026629 | 0.033660 | 0.045438 | 0.040705 | 0.034630 | 0.029781 |
| R10 | 0.023965 | 0.020918 | 0.030485 | 0.086752 | 0.037213 | 0.027152 | 0.023348 |
| RN  | 0.004965 | 0.003339 | 0.012817 | 0.005078 | 0.008598 | 0.013715 | 0.013484 |
| ZU  | 0.310832 | 0.294770 | 0.310186 | 0.304159 | 0.316434 | 0.312512 | 0.308642 |
| ZR  | 0.022502 | 0.038564 | 0.023149 | 0.029176 | 0.016915 | 0.020820 | 0.024691 |

Supplementary Column Profiles

|     | L711     | L721     | L731     | L712     | L722     | L732     | L713     | L723     | L733     |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| I0  | 0.031900 | 0.118242 | 0.063119 | 0.113460 | 0.112374 | 0.202436 | 0.066733 | 0.085515 | 0.153870 |
| I1  | 0.077289 | 0.163846 | 0.218049 | 0.135149 | 0.150054 | 0.116043 | 0.200533 | 0.183758 | 0.163469 |
| I2  | 0.040171 | 0.000000 | 0.000000 | 0.028125 | 0.010367 | 0.002971 | 0.016642 | 0.018251 | 0.003157 |
| I3  | 0.039196 | 0.000000 | 0.000000 | 0.024110 | 0.023066 | 0.003001 | 0.000000 | 0.024446 | 0.004808 |
| I4  | 0.144778 | 0.051246 | 0.052165 | 0.032489 | 0.037472 | 0.008883 | 0.049426 | 0.021364 | 0.006427 |
| IN  | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.001603 |
| R0  | 0.057776 | 0.000000 | 0.000000 | 0.008553 | 0.004122 | 0.059685 | 0.000000 | 0.003021 | 0.004760 |
| R1  | 0.004963 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.003030 | 0.067066 | 0.015351 | 0.014360 |
| R2  | 0.000000 | 0.000000 | 0.000000 | 0.006720 | 0.006308 | 0.026708 | 0.067399 | 0.037020 | 0.035210 |
| R3  | 0.017816 | 0.023507 | 0.000000 | 0.006655 | 0.012449 | 0.006001 | 0.016475 | 0.036745 | 0.022453 |
| R4  | 0.000000 | 0.055242 | 0.000000 | 0.008684 | 0.004143 | 0.017706 | 0.033117 | 0.045810 | 0.024088 |
| R5  | 0.000000 | 0.000000 | 0.113138 | 0.010801 | 0.029124 | 0.036037 | 0.049426 | 0.018251 | 0.036717 |
| R6  | 0.040002 | 0.000000 | 0.000000 | 0.015426 | 0.010513 | 0.044534 | 0.000000 | 0.030611 | 0.032037 |
| R7  | 0.026555 | 0.092149 | 0.052686 | 0.080797 | 0.047943 | 0.041355 | 0.050258 | 0.036684 | 0.033752 |
| R8  | 0.040468 | 0.000000 | 0.063119 | 0.071742 | 0.064514 | 0.032799 | 0.016642 | 0.024385 | 0.040146 |
| R9  | 0.069483 | 0.000000 | 0.052165 | 0.065414 | 0.048110 | 0.029857 | 0.032951 | 0.042849 | 0.033623 |
| R10 | 0.053915 | 0.134932 | 0.052165 | 0.054221 | 0.087351 | 0.023708 | 0.000000 | 0.015260 | 0.027277 |
| RN  | 0.022355 | 0.027504 | 0.000000 | 0.004320 | 0.018757 | 0.011913 | 0.000000 | 0.027345 | 0.028912 |
| ZU  | 0.301179 | 0.262106 | 0.333333 | 0.256573 | 0.253664 | 0.211022 | 0.282909 | 0.287157 | 0.275895 |
| ZR  | 0.032154 | 0.071227 | 0.000000 | 0.076760 | 0.079669 | 0.122311 | 0.050424 | 0.046176 | 0.057439 |

|     | L714     | L724     | L734     | L715     | L725     | L735     | L811     | L821     | L831     |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| I0  | 0.053534 | 0.074377 | 0.152931 | 0.026723 | 0.089153 | 0.195051 | 0.042103 | 0.073896 | 0.093023 |
| I1  | 0.195199 | 0.194110 | 0.151444 | 0.177139 | 0.206231 | 0.120150 | 0.123652 | 0.111223 | 0.127519 |
| I2  | 0.039141 | 0.021496 | 0.010566 | 0.070959 | 0.015473 | 0.002379 | 0.074131 | 0.033788 | 0.045155 |
| I3  | 0.023959 | 0.027730 | 0.007827 | 0.038911 | 0.017701 | 0.003414 | 0.063426 | 0.050640 | 0.007946 |
| I4  | 0.018798 | 0.015621 | 0.010566 | 0.019452 | 0.015828 | 0.012340 | 0.029327 | 0.059403 | 0.059690 |
| IN  | 0.002702 | 0.000000 | 0.000000 | 0.000149 | 0.001777 | 0.000000 | 0.000694 | 0.004382 | 0.000000 |
| R0  | 0.001368 | 0.000000 | 0.007670 | 0.001270 | 0.004167 | 0.001530 | 0.000174 | 0.000000 | 0.000000 |
| R1  | 0.016885 | 0.006735 | 0.007670 | 0.032111 | 0.012162 | 0.046079 | 0.001627 | 0.002275 | 0.032364 |
| R2  | 0.019120 | 0.028447 | 0.029037 | 0.033386 | 0.024356 | 0.044819 | 0.000325 | 0.000000 | 0.029264 |
| R3  | 0.020277 | 0.029378 | 0.032245 | 0.035868 | 0.025664 | 0.034251 | 0.003720 | 0.011038 | 0.015891 |
| R4  | 0.028266 | 0.035182 | 0.007983 | 0.032832 | 0.020528 | 0.029746 | 0.009382 | 0.018537 | 0.038566 |
| R5  | 0.027687 | 0.023932 | 0.033028 | 0.032778 | 0.016232 | 0.018338 | 0.012408 | 0.009858 | 0.024806 |
| R6  | 0.030916 | 0.014617 | 0.007514 | 0.033071 | 0.040070 | 0.018189 | 0.036257 | 0.016346 | 0.012597 |
| R7  | 0.043579 | 0.021210 | 0.018940 | 0.028582 | 0.004264 | 0.034652 | 0.060790 | 0.032693 | 0.011047 |
| R8  | 0.048320 | 0.056606 | 0.060499 | 0.028498 | 0.043672 | 0.023319 | 0.072287 | 0.041877 | 0.029651 |
| R9  | 0.042593 | 0.021066 | 0.043985 | 0.027177 | 0.034515 | 0.025399 | 0.074445 | 0.070273 | 0.050000 |
| R10 | 0.019192 | 0.057896 | 0.031228 | 0.020038 | 0.027780 | 0.024028 | 0.060422 | 0.122514 | 0.089147 |
| RN  | 0.035130 | 0.038263 | 0.053534 | 0.027722 | 0.054267 | 0.032983 | 0.001497 | 0.007920 | 0.000000 |
| ZU  | 0.283357 | 0.284322 | 0.239728 | 0.261953 | 0.302248 | 0.260848 | 0.319180 | 0.309993 | 0.287791 |
| ZR  | 0.049977 | 0.049011 | 0.093606 | 0.071380 | 0.043914 | 0.072485 | 0.014154 | 0.023340 | 0.045543 |

|     | L812     | L822     | L832     | L813     | L823     | L833     | L814     | L824     | L834     |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| I0  | 0.051904 | 0.070281 | 0.174846 | 0.040995 | 0.072132 | 0.157043 | 0.034125 | 0.049768 | 0.148204 |
| I1  | 0.162221 | 0.160441 | 0.120368 | 0.147865 | 0.190905 | 0.140085 | 0.154039 | 0.195686 | 0.141882 |
| I2  | 0.067840 | 0.035919 | 0.015426 | 0.096409 | 0.036451 | 0.019746 | 0.090638 | 0.045196 | 0.024232 |
| I3  | 0.030725 | 0.019323 | 0.008503 | 0.014136 | 0.015915 | 0.009375 | 0.041693 | 0.020591 | 0.010987 |
| I4  | 0.019302 | 0.047369 | 0.014190 | 0.033927 | 0.017377 | 0.007084 | 0.011025 | 0.021150 | 0.008027 |
| IN  | 0.001340 | 0.000000 | 0.000000 | 0.000000 | 0.000553 | 0.000000 | 0.001814 | 0.000942 | 0.000000 |
| R0  | 0.004319 | 0.005064 | 0.007209 | 0.000000 | 0.006279 | 0.000356 | 0.001140 | 0.000000 | 0.002308 |
| R1  | 0.003023 | 0.002087 | 0.005897 | 0.047498 | 0.028829 | 0.036917 | 0.014671 | 0.011901 | 0.034969 |
| R2  | 0.002874 | 0.009655 | 0.082287 | 0.040995 | 0.044350 | 0.026368 | 0.017769 | 0.016403 | 0.029701 |
| R3  | 0.007030 | 0.010267 | 0.017176 | 0.023466 | 0.029599 | 0.029572 | 0.024487 | 0.026594 | 0.023630 |
| R4  | 0.016264 | 0.026654 | 0.042323 | 0.048629 | 0.029481 | 0.029667 | 0.026922 | 0.033574 | 0.040739 |
| R5  | 0.036430 | 0.028602 | 0.020068 | 0.041843 | 0.036767 | 0.035327 | 0.031426 | 0.017625 | 0.026540 |
| R6  | 0.043191 | 0.023956 | 0.027011 | 0.026011 | 0.032818 | 0.035232 | 0.041148 | 0.038565 | 0.032260 |
| R7  | 0.066247 | 0.030870 | 0.027125 | 0.013288 | 0.034082 | 0.031221 | 0.053123 | 0.045301 | 0.031407 |
| R8  | 0.056417 | 0.041136 | 0.027220 | 0.035058 | 0.034872 | 0.033855 | 0.053931 | 0.036750 | 0.019717 |
| R9  | 0.052202 | 0.046617 | 0.024880 | 0.032796 | 0.023261 | 0.038709 | 0.043230 | 0.056469 | 0.027544 |
| R10 | 0.043951 | 0.106020 | 0.047573 | 0.023749 | 0.029659 | 0.031826 | 0.020252 | 0.045859 | 0.063265 |
| RN  | 0.001385 | 0.002407 | 0.004565 | 0.000000 | 0.003337 | 0.004284 | 0.005233 | 0.004293 | 0.001254 |
| ZU  | 0.294938 | 0.289011 | 0.280815 | 0.299406 | 0.301483 | 0.287766 | 0.311908 | 0.315185 | 0.284066 |
| ZR  | 0.038396 | 0.044322 | 0.052518 | 0.033927 | 0.031850 | 0.045568 | 0.021426 | 0.018148 | 0.049268 |

|    | L815     | L825     | L835     | L911     | L921     | L931     | L912     | L922     | L932     |
|----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| I0 | 0.014806 | 0.043912 | 0.178917 | 0.014137 | 0.024874 | 0.123194 | 0.027880 | 0.036405 | 0.127316 |
| I1 | 0.107098 | 0.186751 | 0.135578 | 0.068324 | 0.135453 | 0.170573 | 0.093519 | 0.131368 | 0.157763 |

|     |          |          |          |          |          |          |          |          |          |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| I2  | 0.130271 | 0.055208 | 0.007209 | 0.153489 | 0.079548 | 0.018133 | 0.113126 | 0.050595 | 0.018898 |
| I3  | 0.064114 | 0.028188 | 0.006804 | 0.077876 | 0.053689 | 0.009748 | 0.062937 | 0.044039 | 0.012346 |
| I4  | 0.015449 | 0.017737 | 0.004826 | 0.018958 | 0.039788 | 0.011617 | 0.035723 | 0.070926 | 0.017011 |
| IN  | 0.001594 | 0.001538 | 0.000000 | 0.000548 | 0.000000 | 0.000000 | 0.000139 | 0.000000 | 0.000000 |
| RO  | 0.000567 | 0.000145 | 0.000000 | 0.003409 | 0.004871 | 0.002627 | 0.018873 | 0.010168 | 0.005794 |
| R1  | 0.033805 | 0.019476 | 0.043626 | 0.025974 | 0.018225 | 0.033690 | 0.001801 | 0.004072 | 0.021224 |
| R2  | 0.035428 | 0.022351 | 0.036589 | 0.028203 | 0.020332 | 0.029397 | 0.009741 | 0.015092 | 0.074529 |
| R3  | 0.039846 | 0.031215 | 0.036589 | 0.032824 | 0.030511 | 0.030155 | 0.009090 | 0.010742 | 0.016931 |
| R4  | 0.038429 | 0.034355 | 0.034683 | 0.032956 | 0.030949 | 0.027932 | 0.016559 | 0.012148 | 0.021104 |
| R5  | 0.035997 | 0.036207 | 0.029003 | 0.036157 | 0.026844 | 0.033135 | 0.031663 | 0.024825 | 0.045664 |
| R6  | 0.036296 | 0.036754 | 0.027646 | 0.035195 | 0.030155 | 0.026770 | 0.028974 | 0.017947 | 0.020612 |
| R7  | 0.033102 | 0.041931 | 0.029479 | 0.033487 | 0.032262 | 0.034094 | 0.046767 | 0.028393 | 0.026061 |
| R8  | 0.029662 | 0.037503 | 0.033092 | 0.030377 | 0.038967 | 0.026215 | 0.052503 | 0.033808 | 0.027005 |
| R9  | 0.026331 | 0.038316 | 0.029614 | 0.033117 | 0.037599 | 0.039954 | 0.060568 | 0.047361 | 0.026872 |
| R10 | 0.020519 | 0.031360 | 0.030720 | 0.028777 | 0.050104 | 0.036973 | 0.051866 | 0.122826 | 0.044162 |
| RN  | 0.003350 | 0.003720 | 0.002292 | 0.012856 | 0.012478 | 0.012476 | 0.004933 | 0.005951 | 0.003376 |
| ZU  | 0.294450 | 0.313342 | 0.291235 | 0.309357 | 0.320518 | 0.311850 | 0.301481 | 0.309158 | 0.296189 |
| ZR  | 0.038883 | 0.019991 | 0.042098 | 0.023978 | 0.012834 | 0.021467 | 0.031857 | 0.024176 | 0.037145 |

|     |          |          |          |          |          |          |          |          |          |
|-----|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|     | L913     | L923     | L933     | L914     | L924     | L934     | L915     | L925     | L935     |
| I0  | 0.019950 | 0.036991 | 0.107378 | 0.012990 | 0.031100 | 0.131686 | 0.014067 | 0.028371 | 0.135606 |
| I1  | 0.085845 | 0.155128 | 0.175318 | 0.063878 | 0.129612 | 0.164359 | 0.072923 | 0.148752 | 0.161897 |
| I2  | 0.130981 | 0.062484 | 0.023036 | 0.160151 | 0.083573 | 0.011223 | 0.157992 | 0.078132 | 0.016563 |
| I3  | 0.071847 | 0.037157 | 0.016454 | 0.077360 | 0.041786 | 0.014300 | 0.071755 | 0.045123 | 0.009754 |
| I4  | 0.023711 | 0.041573 | 0.011146 | 0.018371 | 0.047262 | 0.011766 | 0.016090 | 0.032814 | 0.009513 |
| IN  | 0.001000 | 0.000000 | 0.000000 | 0.000586 | 0.000000 | 0.000000 | 0.000507 | 0.000142 | 0.000000 |
| RO  | 0.012094 | 0.006415 | 0.001486 | 0.003067 | 0.003296 | 0.000272 | 0.002705 | 0.002458 | 0.000698 |
| R1  | 0.009475 | 0.028743 | 0.029883 | 0.022222 | 0.022913 | 0.028147 | 0.027694 | 0.023199 | 0.033663 |
| R2  | 0.018474 | 0.027243 | 0.026168 | 0.026168 | 0.028868 | 0.029233 | 0.029952 | 0.023332 | 0.032840 |
| R3  | 0.031091 | 0.019620 | 0.032219 | 0.033642 | 0.022222 | 0.028871 | 0.035412 | 0.028537 | 0.032898 |
| R4  | 0.026996 | 0.029784 | 0.031423 | 0.032221 | 0.023020 | 0.027242 | 0.035708 | 0.031221 | 0.030418 |
| R5  | 0.030615 | 0.037366 | 0.029512 | 0.038058 | 0.032483 | 0.030229 | 0.036424 | 0.035047 | 0.031492 |
| R6  | 0.028567 | 0.023244 | 0.026645 | 0.033689 | 0.033705 | 0.022626 | 0.033944 | 0.032651 | 0.029540 |
| R7  | 0.044184 | 0.032159 | 0.029087 | 0.034044 | 0.035354 | 0.038918 | 0.034463 | 0.034105 | 0.030942 |
| R8  | 0.054945 | 0.034158 | 0.039915 | 0.037215 | 0.031898 | 0.030953 | 0.032057 | 0.036498 | 0.033592 |
| R9  | 0.045470 | 0.037949 | 0.038907 | 0.033722 | 0.044923 | 0.039823 | 0.029139 | 0.037495 | 0.033901 |
| R10 | 0.023140 | 0.048405 | 0.038641 | 0.025614 | 0.041733 | 0.040818 | 0.022237 | 0.035143 | 0.033399 |
| RN  | 0.008285 | 0.008248 | 0.009395 | 0.013670 | 0.012919 | 0.016201 | 0.013596 | 0.013647 | 0.009951 |
| ZU  | 0.320383 | 0.317171 | 0.311093 | 0.312198 | 0.318660 | 0.309892 | 0.308138 | 0.318292 | 0.305199 |
| ZR  | 0.012951 | 0.016163 | 0.022293 | 0.021134 | 0.014673 | 0.023441 | 0.025196 | 0.015041 | 0.028135 |

Inertia and Chi-Square Decomposition

| Singular Values | Principal Inertias | Chi-Squares                | Percents | 12    | 24    | 36    | 48    | 60    |
|-----------------|--------------------|----------------------------|----------|-------|-------|-------|-------|-------|
| 0.20777         | 0.04317            | 972901                     | 58.65%   | ***** | ***** | ***** | ***** | ***** |
| 0.13193         | 0.01741            | 351964                     | 23.65%   | ***** | ***** | ***** | ***** | ***** |
| 0.07313         | 0.00535            | 108147                     | 7.27%    | ***   | ***   | ***   | ***   | ***   |
| 0.05732         | 0.00329            | 66444.7                    | 4.46%    | **    | **    | **    | **    | **    |
| 0.04788         | 0.00229            | 46362.9                    | 3.12%    | *     | *     | *     | *     | *     |
| 0.03251         | 0.00106            | 21372.4                    | 1.44%    | *     | *     | *     | *     | *     |
| 0.02487         | 0.00062            | 12507.8                    | 0.84%    |       |       |       |       |       |
| 0.01366         | 0.00019            | 3770.96                    | 0.25%    |       |       |       |       |       |
| 0.01072         | 0.00012            | 2325.67                    | 0.16%    |       |       |       |       |       |
| 0.00818         | 0.00007            | 1352.87                    | 0.09%    |       |       |       |       |       |
| 0.00530         | 0.00003            | 567.967                    | 0.04%    |       |       |       |       |       |
| 0.00336         | 0.00001            | 228.055                    | 0.02%    |       |       |       |       |       |
| 0.00309         | 0.00001            | 193.129                    | 0.01%    |       |       |       |       |       |
| 0.00247         | 0.00001            | 123.226                    | 0.01%    |       |       |       |       |       |
| 0.07360         | 1488263            | (Degrees of Freedom = 266) |          |       |       |       |       |       |

Row Coordinates

|     | Dim1     | Dim2     |
|-----|----------|----------|
| I0  | 2.43849  | 2.92342  |
| I1  | 1.48253  | -0.74403 |
| I2  | -1.48844 | -0.11473 |
| I3  | -1.21906 | -0.05221 |
| I4  | 0.54536  | 1.50175  |
| IN  | -1.20596 | -0.46742 |
| RO  | 1.16215  | 7.52925  |
| R1  | -0.19239 | -1.44249 |
| R2  | -0.02296 | -0.77071 |
| R3  | -0.33007 | -1.06143 |
| R4  | -0.37239 | -0.68323 |
| R5  | -0.29478 | -0.20596 |
| R6  | -0.17366 | -0.27480 |
| R7  | -0.10240 | 0.69953  |
| R8  | 0.10235  | 0.90361  |
| R9  | 0.18534  | 1.16779  |
| R10 | 0.43802  | 2.76246  |
| RN  | 2.11129  | -1.87408 |

|    |          |          |
|----|----------|----------|
| ZU | -0.25608 | 0.20649  |
| ZR | 1.88579  | -1.52254 |

Summary Statistics for the Row Points

|     | Quality  | Mass     | Inertia  |
|-----|----------|----------|----------|
| I0  | 0.857804 | 0.025845 | 0.165976 |
| I1  | 0.959590 | 0.115500 | 0.170922 |
| I2  | 0.980483 | 0.116446 | 0.154693 |
| I3  | 0.970399 | 0.056971 | 0.051211 |
| I4  | 0.418509 | 0.017833 | 0.030160 |
| IN  | 0.111094 | 0.000777 | 0.006331 |
| R0  | 0.579352 | 0.002028 | 0.049694 |
| R1  | 0.808726 | 0.029096 | 0.018485 |
| R2  | 0.482895 | 0.031475 | 0.009176 |
| R3  | 0.862707 | 0.035072 | 0.013429 |
| R4  | 0.796054 | 0.034483 | 0.008305 |
| R5  | 0.636764 | 0.034606 | 0.003315 |
| R6  | 0.441124 | 0.034045 | 0.002744 |
| R7  | 0.488787 | 0.033390 | 0.008326 |
| R8  | 0.536143 | 0.031928 | 0.011865 |
| R9  | 0.743359 | 0.029674 | 0.013678 |
| R10 | 0.770040 | 0.024190 | 0.060229 |
| RN  | 0.497750 | 0.013267 | 0.091827 |
| ZU  | 0.921927 | 0.293390 | 0.015449 |
| ZR  | 0.922333 | 0.039983 | 0.114185 |

Partial Contributions to Inertia for the Row Points

|     | Dim1     | Dim2     |
|-----|----------|----------|
| I0  | 0.153681 | 0.220882 |
| I1  | 0.253857 | 0.063939 |
| I2  | 0.257980 | 0.001533 |
| I3  | 0.084666 | 0.000155 |
| I4  | 0.005304 | 0.040218 |
| IN  | 0.001131 | 0.000170 |
| R0  | 0.002739 | 0.114946 |
| R1  | 0.001077 | 0.060541 |
| R2  | 0.000017 | 0.018696 |
| R3  | 0.003821 | 0.039513 |
| R4  | 0.004782 | 0.016097 |
| R5  | 0.003007 | 0.001468 |
| R6  | 0.001027 | 0.002571 |
| R7  | 0.000350 | 0.016340 |
| R8  | 0.000334 | 0.026070 |
| R9  | 0.001019 | 0.040467 |
| R10 | 0.004641 | 0.184601 |
| RN  | 0.059140 | 0.046597 |
| ZU  | 0.019239 | 0.012510 |
| ZR  | 0.142189 | 0.092686 |

Indices of the Coordinates that Contribute Most to Inertia for the Row Points

|     | Dim1 | Dim2 | Best |
|-----|------|------|------|
| I0  | 2    | 2    | 2    |
| I1  | 1    | 1    | 1    |
| I2  | 1    | 0    | 1    |
| I3  | 0    | 0    | 1    |
| I4  | 0    | 0    | 2    |
| IN  | 0    | 0    | 1    |
| R0  | 0    | 2    | 2    |
| R1  | 0    | 2    | 2    |
| R2  | 0    | 0    | 2    |
| R3  | 0    | 0    | 2    |
| R4  | 0    | 0    | 2    |
| R5  | 0    | 0    | 1    |
| R6  | 0    | 0    | 2    |
| R7  | 0    | 0    | 2    |
| R8  | 0    | 0    | 2    |
| R9  | 0    | 2    | 2    |
| R10 | 0    | 2    | 2    |
| RN  | 0    | 1    | 1    |
| ZU  | 0    | 0    | 1    |
| ZR  | 1    | 1    | 1    |

Squared Cosines for the Row Points

|    | Dim1     | Dim2     |
|----|----------|----------|
| I0 | 0.543076 | 0.314728 |
| I1 | 0.871121 | 0.088469 |
| I2 | 0.978140 | 0.002343 |
| I3 | 0.969682 | 0.000717 |

|     |          |          |
|-----|----------|----------|
| I4  | 0.103147 | 0.315362 |
| IN  | 0.104749 | 0.006345 |
| R0  | 0.032322 | 0.547030 |
| R1  | 0.034170 | 0.774555 |
| R2  | 0.001060 | 0.481835 |
| R3  | 0.166882 | 0.695825 |
| R4  | 0.337639 | 0.458355 |
| R5  | 0.532039 | 0.104725 |
| R6  | 0.219504 | 0.221621 |
| R7  | 0.024666 | 0.464121 |
| R8  | 0.016533 | 0.519611 |
| R9  | 0.043706 | 0.699654 |
| R10 | 0.045197 | 0.724844 |
| RN  | 0.377742 | 0.120008 |
| SU  | 0.730426 | 0.191500 |
| SR  | 0.730367 | 0.191966 |

## Column Coordinates

|     | Dim1      | Dim2      |
|-----|-----------|-----------|
| L71 | 0.354496  | 0.824205  |
| L72 | 0.678312  | 0.659366  |
| L73 | 0.610488  | 0.255334  |
| L74 | 0.454618  | 0.072360  |
| L75 | 0.296513  | -0.130912 |
| L81 | 0.125058  | 0.468614  |
| L82 | 0.406989  | 0.458436  |
| L83 | 0.472572  | 0.298989  |
| L84 | 0.116754  | 0.136215  |
| L85 | -0.084193 | -0.045318 |
| L91 | -0.164771 | 0.081095  |
| L92 | 0.244024  | 0.507620  |
| L93 | 0.147970  | 0.253272  |
| L94 | -0.185236 | 0.085318  |
| L95 | -0.164080 | 0.032195  |

## Supplementary Column Coordinates

|      | Dim1      | Dim2      |
|------|-----------|-----------|
| L711 | 0.285739  | 0.901023  |
| L721 | 0.705789  | 0.569365  |
| L731 | 0.420485  | 0.445055  |
| L712 | 0.545106  | 0.590733  |
| L722 | 0.635707  | 0.557240  |
| L732 | 0.920485  | 0.898558  |
| L713 | 0.440726  | -0.010460 |
| L723 | 0.480107  | 0.109222  |
| L733 | 0.695303  | 0.357657  |
| L714 | 0.419247  | 0.042151  |
| L724 | 0.510806  | 0.159652  |
| L734 | 0.814291  | 0.336613  |
| L715 | 0.280073  | -0.147246 |
| L725 | 0.601694  | 0.120208  |
| L735 | 0.767321  | 0.366993  |
| L811 | 0.088243  | 0.442062  |
| L821 | 0.292426  | 0.674577  |
| L831 | 0.397945  | 0.469346  |
| L812 | 0.243019  | 0.317552  |
| L822 | 0.408189  | 0.531951  |
| L832 | 0.614765  | 0.537849  |
| L813 | 0.130183  | 0.030019  |
| L823 | 0.364804  | 0.164970  |
| L833 | 0.551707  | 0.390818  |
| L814 | 0.081789  | 0.114142  |
| L824 | 0.285905  | 0.216683  |
| L834 | 0.539962  | 0.441208  |
| L815 | -0.103198 | -0.057264 |
| L825 | 0.216677  | 0.111035  |
| L835 | 0.602280  | 0.422723  |
| L911 | -0.207948 | 0.061343  |
| L921 | 0.064326  | 0.210482  |
| L931 | 0.491631  | 0.336005  |
| L912 | 0.006485  | 0.444457  |
| L922 | 0.227516  | 0.597300  |
| L932 | 0.506643  | 0.379951  |
| L913 | -0.137865 | 0.236851  |
| L923 | 0.168347  | 0.233448  |
| L933 | 0.440657  | 0.296838  |
| L914 | -0.231918 | 0.065648  |
| L924 | 0.085452  | 0.208489  |
| L934 | 0.522272  | 0.367926  |
| L915 | -0.204344 | 0.015756  |
| L925 | 0.095201  | 0.124361  |
| L935 | 0.513412  | 0.334324  |

## Summary Statistics for the Column Points

|     | Quality  | Mass     | Inertia  |
|-----|----------|----------|----------|
| L71 | 0.424677 | 0.001471 | 0.037884 |
| L72 | 0.872718 | 0.006307 | 0.087863 |
| L73 | 0.743210 | 0.005004 | 0.040053 |
| L74 | 0.808417 | 0.008844 | 0.031497 |
| L75 | 0.980082 | 0.217055 | 0.316108 |
| L81 | 0.613292 | 0.005402 | 0.028152 |
| L82 | 0.918121 | 0.009476 | 0.052695 |
| L83 | 0.626405 | 0.006847 | 0.046442 |
| L84 | 0.319669 | 0.021190 | 0.028986 |
| L85 | 0.541518 | 0.307649 | 0.070566 |
| L91 | 0.856814 | 0.027261 | 0.014579 |
| L92 | 0.800250 | 0.015136 | 0.081569 |
| L93 | 0.809634 | 0.003158 | 0.004559 |
| L94 | 0.846063 | 0.015153 | 0.010120 |
| L95 | 0.892860 | 0.350049 | 0.148925 |

## Quality of Representation for the Supplementary Column Points

|      |          |
|------|----------|
| L711 | 0.326278 |
| L721 | 0.567150 |
| L731 | 0.426250 |
| L712 | 0.829044 |
| L722 | 0.880026 |
| L732 | 0.506422 |
| L713 | 0.369347 |
| L723 | 0.738964 |
| L733 | 0.702795 |
| L714 | 0.751678 |
| L724 | 0.724492 |
| L734 | 0.708233 |
| L715 | 0.964941 |
| L725 | 0.705555 |
| L735 | 0.529536 |
| L811 | 0.537327 |
| L821 | 0.643696 |
| L831 | 0.622628 |
| L812 | 0.615711 |
| L822 | 0.760730 |
| L832 | 0.576542 |
| L813 | 0.142155 |
| L823 | 0.654531 |
| L833 | 0.559615 |
| L814 | 0.201718 |
| L824 | 0.527648 |
| L834 | 0.614561 |
| L815 | 0.657018 |
| L825 | 0.428351 |
| L835 | 0.496385 |
| L911 | 0.858751 |
| L921 | 0.437343 |
| L931 | 0.652130 |
| L912 | 0.611301 |
| L922 | 0.556062 |
| L932 | 0.617486 |
| L913 | 0.537166 |
| L923 | 0.557436 |
| L933 | 0.666998 |
| L914 | 0.848543 |
| L924 | 0.440251 |
| L934 | 0.661518 |
| L915 | 0.895282 |
| L925 | 0.358863 |
| L935 | 0.599071 |

## Partial Contributions to Inertia for the Column Points

|     | Dim1     | Dim2     |
|-----|----------|----------|
| L71 | 0.004282 | 0.057410 |
| L72 | 0.067220 | 0.157528 |
| L73 | 0.043197 | 0.018740 |
| L74 | 0.042340 | 0.002660 |
| L75 | 0.442050 | 0.213701 |
| L81 | 0.001957 | 0.068151 |
| L82 | 0.036357 | 0.114406 |
| L83 | 0.035421 | 0.035164 |
| L84 | 0.006691 | 0.022587 |
| L85 | 0.050516 | 0.036298 |
| L91 | 0.017144 | 0.010299 |
| L92 | 0.020878 | 0.224236 |
| L93 | 0.001602 | 0.011637 |

L94 0.012043 0.006337  
 L95 0.218303 0.020845

Indices of the Coordinates that Contribute Most to Inertia for the Column Points

|     | Dim1 | Dim2 | Best |
|-----|------|------|------|
| L71 | 0    | 2    | 2    |
| L72 | 2    | 2    | 2    |
| L73 | 1    | 0    | 1    |
| L74 | 0    | 0    | 1    |
| L75 | 1    | 1    | 1    |
| L81 | 0    | 2    | 2    |
| L82 | 0    | 2    | 2    |
| L83 | 0    | 0    | 1    |
| L84 | 0    | 0    | 2    |
| L85 | 1    | 0    | 1    |
| L91 | 0    | 0    | 1    |
| L92 | 0    | 2    | 2    |
| L93 | 0    | 0    | 2    |
| L94 | 0    | 0    | 1    |
| L95 | 1    | 0    | 1    |

Squared Cosines for the Column Points

|     | Dim1     | Dim2     |
|-----|----------|----------|
| L71 | 0.066297 | 0.358380 |
| L72 | 0.448717 | 0.424001 |
| L73 | 0.632557 | 0.110653 |
| L74 | 0.788442 | 0.019975 |
| L75 | 0.820203 | 0.159879 |
| L81 | 0.040774 | 0.572518 |
| L82 | 0.404673 | 0.513448 |
| L83 | 0.447340 | 0.179065 |
| L84 | 0.135386 | 0.184282 |
| L85 | 0.419869 | 0.121649 |
| L91 | 0.689738 | 0.167076 |
| L92 | 0.150122 | 0.650128 |
| L93 | 0.206029 | 0.603606 |
| L94 | 0.697988 | 0.148076 |
| L95 | 0.859758 | 0.033102 |

Squared Cosines for the Supplementary Column Points

|      | Dim1     | Dim2     |
|------|----------|----------|
| L711 | 0.029815 | 0.296463 |
| L721 | 0.343566 | 0.223584 |
| L731 | 0.201034 | 0.225215 |
| L712 | 0.381272 | 0.447772 |
| L722 | 0.497648 | 0.382378 |
| L732 | 0.259314 | 0.247107 |
| L713 | 0.369139 | 0.000208 |
| L723 | 0.702601 | 0.036363 |
| L733 | 0.555746 | 0.147049 |
| L714 | 0.744155 | 0.007522 |
| L724 | 0.660017 | 0.064475 |
| L734 | 0.604870 | 0.103363 |
| L715 | 0.755984 | 0.208957 |
| L725 | 0.678475 | 0.027080 |
| L735 | 0.430955 | 0.098581 |
| L811 | 0.020590 | 0.516736 |
| L821 | 0.101827 | 0.541868 |
| L831 | 0.260400 | 0.362228 |
| L812 | 0.227414 | 0.388297 |
| L822 | 0.281927 | 0.478803 |
| L832 | 0.326574 | 0.249968 |
| L813 | 0.134978 | 0.007177 |
| L823 | 0.543405 | 0.111126 |
| L833 | 0.372629 | 0.186986 |
| L814 | 0.068434 | 0.133283 |
| L824 | 0.335144 | 0.192504 |
| L834 | 0.368515 | 0.246046 |
| L815 | 0.502344 | 0.154675 |
| L825 | 0.339261 | 0.089091 |
| L835 | 0.332559 | 0.163826 |
| L911 | 0.790006 | 0.068745 |
| L921 | 0.037358 | 0.399984 |
| L931 | 0.444502 | 0.207628 |
| L912 | 0.000130 | 0.611171 |
| L922 | 0.070457 | 0.485605 |
| L932 | 0.395215 | 0.222272 |
| L913 | 0.135940 | 0.401226 |
| L923 | 0.190711 | 0.366726 |
| L933 | 0.458805 | 0.208193 |

| OBS  | TYPE | NAME | QUANT | MARS | INERTIA | D1       | D2       | CORR    |         | SQU     |         | BB  |     |
|------|------|------|-------|------|---------|----------|----------|---------|---------|---------|---------|-----|-----|
|      |      |      |       |      |         |          |          | 1       | 2       | 1       | 2       | 1   | 2   |
| L914 |      |      |       |      | 0.07360 |          |          | 0.04317 | 0.01741 |         |         |     |     |
| L924 |      |      |       |      | 0.16598 | 2.43849  | 2.92342  | 0.15368 | 0.22088 | 0.54308 | 0.31473 | 2.2 | 2.2 |
| L934 |      |      |       |      | 0.17092 | 1.48253  | -0.74403 | 0.25386 | 0.36394 | 0.87112 | 0.08847 | 1.1 | 1.1 |
| L915 |      |      |       |      | 0.00633 | -1.48844 | -0.11473 | 0.25798 | 0.00153 | 0.97814 | 0.00234 | 1.0 | 1.0 |
| L925 |      |      |       |      | 0.05121 | -1.21906 | -0.05221 | 0.08467 | 0.00016 | 0.96968 | 0.00072 | 0.0 | 0.1 |
| L935 |      |      |       |      | 0.03016 | 0.54536  | 1.50175  | 0.00530 | 0.04022 | 0.10315 | 0.31536 | 0.0 | 0.2 |
|      |      |      |       |      | 0.00078 | -1.20596 | -0.46742 | 0.00113 | 0.00017 | 0.10475 | 0.00634 | 0.0 | 0.1 |
|      |      |      |       |      | 0.04969 | 1.16215  | 7.52925  | 0.00274 | 0.11495 | 0.03232 | 0.54703 | 0.2 | 2.2 |
|      |      |      |       |      | 0.01848 | -0.19239 | -1.44249 | 0.00108 | 0.06054 | 0.03417 | 0.77456 | 0.2 | 2.2 |
|      |      |      |       |      | 0.00918 | -0.02296 | -0.77071 | 0.00002 | 0.01870 | 0.00106 | 0.48183 | 0.0 | 0.2 |
|      |      |      |       |      | 0.01343 | -0.33007 | -1.06143 | 0.00382 | 0.03951 | 0.16688 | 0.69582 | 0.0 | 0.2 |
|      |      |      |       |      | 0.00831 | -0.37239 | -0.68323 | 0.00478 | 0.01610 | 0.33770 | 0.45836 | 0.0 | 0.2 |
|      |      |      |       |      | 0.00332 | -0.29478 | -0.20596 | 0.00301 | 0.00147 | 0.53204 | 0.10472 | 0.0 | 0.1 |
|      |      |      |       |      | 0.00274 | -0.17366 | -0.27480 | 0.00103 | 0.00257 | 0.21950 | 0.22162 | 0.0 | 0.2 |
|      |      |      |       |      | 0.00833 | -0.10240 | 0.69953  | 0.00035 | 0.01634 | 0.02467 | 0.46412 | 0.0 | 0.2 |
|      |      |      |       |      | 0.01187 | 0.10235  | 0.90361  | 0.00033 | 0.02607 | 0.01653 | 0.51961 | 0.0 | 0.2 |
|      |      |      |       |      | 0.01368 | 0.18534  | 1.16779  | 0.00102 | 0.64047 | 0.04371 | 0.69965 | 0.2 | 2.2 |
|      |      |      |       |      | 0.06023 | 0.43802  | 2.76246  | 0.00464 | 0.18460 | 0.04520 | 0.72484 | 0.2 | 2.2 |
|      |      |      |       |      | 0.09183 | 2.11129  | -1.87408 | 0.05914 | 0.04660 | 0.37774 | 0.12001 | 0.1 | 1.1 |
|      |      |      |       |      | 0.01545 | -0.25608 | 0.20649  | 0.01924 | 0.01251 | 0.73043 | 0.19150 | 0.0 | 0.1 |
|      |      |      |       |      | 0.11419 | 1.88579  | -1.52254 | 0.14219 | 0.09269 | 0.73037 | 0.19197 | 1.1 | 1.1 |
|      |      |      |       |      | 0.03788 | 0.35450  | 0.82420  | 0.00428 | 0.05741 | 0.06630 | 0.35838 | 0.2 | 2.2 |
|      |      |      |       |      | 0.00876 | 0.67831  | 0.65937  | 0.06722 | 0.15753 | 0.44872 | 0.42400 | 2.2 | 2.2 |
|      |      |      |       |      | 0.04005 | 0.61049  | 0.25533  | 0.04320 | 0.01874 | 0.63256 | 0.11065 | 1.0 | 1.0 |
|      |      |      |       |      | 0.03150 | 0.45462  | 0.07236  | 0.04234 | 0.00266 | 0.78844 | 0.01997 | 0.0 | 0.1 |
|      |      |      |       |      | 0.31611 | 0.29651  | -0.13091 | 0.44205 | 0.21370 | 0.82020 | 0.15988 | 1.1 | 1.1 |
|      |      |      |       |      | 0.02815 | 0.12506  | 0.46861  | 0.00196 | 0.06815 | 0.04077 | 0.57252 | 0.2 | 2.2 |
|      |      |      |       |      | 0.05270 | 0.40699  | 0.45844  | 0.03636 | 0.11441 | 0.40467 | 0.51345 | 0.2 | 2.2 |
|      |      |      |       |      | 0.04644 | 0.47257  | 0.29899  | 0.03542 | 0.03516 | 0.44734 | 0.17907 | 0.0 | 0.1 |
|      |      |      |       |      | 0.02899 | 0.11675  | 0.13622  | 0.00669 | 0.02259 | 0.13539 | 0.18428 | 0.0 | 0.2 |
|      |      |      |       |      | 0.07057 | -0.08419 | -0.04532 | 0.05052 | 0.03630 | 0.41987 | 0.12165 | 1.0 | 1.0 |
|      |      |      |       |      | 0.01458 | -0.16477 | 0.08110  | 0.01714 | 0.01030 | 0.68974 | 0.16708 | 0.0 | 0.1 |
|      |      |      |       |      | 0.08157 | 0.24402  | 0.50782  | 0.02088 | 0.22424 | 0.15012 | 0.65013 | 0.2 | 2.2 |
|      |      |      |       |      | 0.00456 | 0.14797  | 0.25327  | 0.00160 | 0.01164 | 0.20603 | 0.60361 | 0.0 | 0.2 |
|      |      |      |       |      | 0.01012 | -0.18524 | 0.08532  | 0.01204 | 0.00634 | 0.69799 | 0.14808 | 0.0 | 0.1 |
|      |      |      |       |      | 0.14893 | -0.16408 | 0.03220  | 0.21830 | 0.02084 | 0.85976 | 0.03310 | 1.0 | 1.0 |
|      |      |      |       |      |         | 0.28574  | 0.90102  |         |         | 0.02982 | 0.29646 |     |     |
|      |      |      |       |      |         | 0.70579  | 0.56936  |         |         | 0.34357 | 0.22358 |     |     |
|      |      |      |       |      |         | 0.42048  | 0.44506  |         |         | 0.20103 | 0.22522 |     |     |
|      |      |      |       |      |         | 0.54511  | 0.59073  |         |         | 0.38127 | 0.44777 |     |     |
|      |      |      |       |      |         | 0.63571  | 0.55724  |         |         | 0.49765 | 0.38238 |     |     |
|      |      |      |       |      |         | 0.92048  | 0.89856  |         |         | 0.25931 | 0.24711 |     |     |
|      |      |      |       |      |         | 0.44073  | -0.01046 |         |         | 0.36914 | 0.00021 |     |     |
|      |      |      |       |      |         | 0.48011  | 0.10922  |         |         | 0.70260 | 0.03636 |     |     |
|      |      |      |       |      |         | 0.69530  | 0.35766  |         |         | 0.55575 | 0.14705 |     |     |
|      |      |      |       |      |         | 0.41925  | 0.04215  |         |         | 0.74416 | 0.00752 |     |     |
|      |      |      |       |      | 0.72449 | 0.51081  | 0.15965  |         | 0.66002 | 0.06447 |         |     |     |
|      |      |      |       |      | 0.70823 | 0.81429  | 0.33661  |         | 0.60487 | 0.10336 |         |     |     |
|      |      |      |       |      | 0.96494 | 0.28007  | -0.14725 |         | 0.75598 | 0.20896 |         |     |     |
|      |      |      |       |      | 0.70555 | 0.60169  | 0.12021  |         | 0.67847 | 0.02708 |         |     |     |
|      |      |      |       |      | 0.52954 | 0.76732  | 0.36699  |         | 0.43096 | 0.09858 |         |     |     |
|      |      |      |       |      | 0.53733 | 0.08824  | 0.44206  |         | 0.02059 | 0.51674 |         |     |     |
|      |      |      |       |      | 0.64370 | 0.29243  | 0.67458  |         | 0.10183 | 0.54187 |         |     |     |
|      |      |      |       |      | 0.62263 | 0.39794  | 0.46935  |         | 0.26040 | 0.36223 |         |     |     |
|      |      |      |       |      | 0.61571 | 0.24302  | 0.31755  |         | 0.22741 | 0.38830 |         |     |     |
|      |      |      |       |      | 0.76073 | 0.40819  | 0.53195  |         | 0.28193 | 0.47880 |         |     |     |
|      |      |      |       |      | 0.57654 | 0.61476  | 0.53785  |         | 0.32657 | 0.24997 |         |     |     |
|      |      |      |       |      | 0.14216 | 0.13018  | 0.03002  |         | 0.13498 | 0.00718 |         |     |     |
|      |      |      |       |      | 0.65453 | 0.36480  | 0.16497  |         | 0.54341 | 0.11113 |         |     |     |
|      |      |      |       |      | 0.55961 | 0.55171  | 0.39082  |         | 0.37263 | 0.18699 |         |     |     |
|      |      |      |       |      | 0.20172 | 0.08179  | 0.11414  |         | 0.06843 | 0.13328 |         |     |     |
|      |      |      |       |      | 0.52765 | 0.28591  | 0.21668  |         | 0.33514 | 0.19250 |         |     |     |
|      |      |      |       |      | 0.61456 | 0.53996  | 0.44121  |         | 0.36852 | 0.24605 |         |     |     |
|      |      |      |       |      | 0.65702 | -0.10320 | -0.05726 |         | 0.50234 | 0.15467 |         |     |     |
|      |      |      |       |      | 0.42835 | 0.21668  | 0.11104  |         | 0.33926 | 0.08909 |         |     |     |
|      |      |      |       |      | 0.49638 | 0.60228  | 0.42272  |         | 0.33256 | 0.16383 |         |     |     |
|      |      |      |       |      | 0.85875 | -0.20795 | 0.06134  |         | 0.79001 | 0.06875 |         |     |     |
|      |      |      |       |      | 0.43734 | 0.06433  | 0.21048  |         | 0.03736 | 0.39998 |         |     |     |
|      |      |      |       |      | 0.65213 | 0.49163  | 0.33600  |         | 0.44450 | 0.20763 |         |     |     |
|      |      |      |       |      | 0.61130 | 0.00648  | 0.44446  |         | 0.00013 | 0.61117 |         |     |     |
|      |      |      |       |      | 0.55606 | 0.22752  | 0.59730  |         | 0.07046 | 0.48560 |         |     |     |
|      |      |      |       |      | 0.61749 | 0.50664  | 0.37995  |         | 0.39521 | 0.22227 |         |     |     |
|      |      |      |       |      | 0.53717 | -0.13787 | 0.23685  |         | 0.13594 | 0.40123 |         |     |     |

---

|    |        |      |         |     |          |         |     |         |         |       |
|----|--------|------|---------|-----|----------|---------|-----|---------|---------|-------|
| 74 | SUPVAR | L923 | 0.55744 | . . | 0.16835  | 0.23345 | . . | 0.19071 | 0.36673 | . . . |
| 75 | SUPVAR | L933 | 0.66700 | . . | 0.44066  | 0.29684 | . . | 0.45881 | 0.20819 | . . . |
| 76 | SUPVAR | L914 | 0.84854 | . . | -0.23192 | 0.06565 | . . | 0.78560 | 0.06295 | . . . |
| 77 | SUPVAR | L924 | 0.44025 | . . | 0.08545  | 0.20849 | . . | 0.06332 | 0.37693 | . . . |
| 78 | SUPVAR | L934 | 0.66152 | . . | 0.52227  | 0.36793 | . . | 0.44211 | 0.21941 | . . . |
| 79 | SUPVAR | L915 | 0.89528 | . . | -0.20434 | 0.01576 | . . | 0.88999 | 0.00529 | . . . |
| 80 | SUPVAR | L925 | 0.35886 | . . | 0.09520  | 0.12436 | . . | 0.13260 | 0.22627 | . . . |
| 81 | SUPVAR | L935 | 0.59907 | . . | 0.51341  | 0.33432 | . . | 0.42069 | 0.17839 | . . . |

## BIBLIOGRAFIA

**BARIONI Jr., W. (1995)** *Análise de Correspondência na Identificação de Fatores de Risco Associado à Diarréia de Leitões na Fase de Lactação*. Piracicaba. ESALQ/USP. Dissertação de Mestrado

**BENZÉCRI, J.-P. (1992)** *Correspondence Analysis Handbook*. Marcel Dekker. New York.

**BOLDRINI, J. L. et all (1980)** *Álgebra Linear* 3ª ed. Harper& Row do Brasil. São Paulo.

**CARDOSO, R. H. A. (1991)** *Análise de Dados Multivariado Através de Técnicas Baseadas nas Decomposição em Valores Singulares*. Campinas. IMECC/UNICAMP. Dissertação de mestrado.

**GREENACRE, M. J. (1984)** *Theory and Applications of Correspondence Analysis*. Academic Press. London.

\_\_\_\_\_ (1988) Correspondence Analysis of multivariate categorical Data by Weighted Least-squares. *Biometrika* 75, nº 3, pp 457-467.

\_\_\_\_\_ (1993) *Correspondence Analysis in Praticce*. Academic Press. San Diego.

**GREENACRE, M. J. and BLASIUS, J. (1994)** *Correspondence Analysis in the Social Sciences*. Academic Press. San Diego.

**GREENACRE, M. J. and HASTIE, T. (1987)** *The Geometric Interpretation of Correspondence Analysis*. *Journal of the Ammerican Statistical Association (JASA)* 82, nº 398, pp 437-447.

## BIBLIOGRAFIA

---

- GRENACRE, M. J. and VRBA, E. S. (1984)** Graphical Display and Interpretation of Antelope Census Data in African Wildlife Areas, using Correspondence Analysis. *Ecology* **65** n° 3 pp. 984-997.
- HALMOS, P.R. (1974)** *Finite-dimensional Vector Spaces*. Springer-Verlag. New York.
- HILL, M. O. (1974)** Correspondence Analysis: a neglected multivariate method. *Applied Statistics*, **23** (3), 340-354.
- HOFFMAN, D. L. and FRANKE, G. R. (1986)** Correspondence Analysis: Grafical Representation of Categorical data in Marketing Research. *Journal of Marketing Research* **23**, pp 213-227.
- HOFFMAN, K e KUNZE, R.(1971)** *Álgebra Linear*. Editora Polígono São Paulo.
- HUBERT, L. and ARABIE, P. Principal (1992)** Correspondence Analysis and Optimal Strutural Representation. *Psychometrika* 58. n° 4 pp. 984-997
- IEMMA, A. F. (1988)** *Matrizes para Estatístico Um Texto para Profissionais de Ciências Aplicadas*. ESALQ/USP. Piracicaba.
- JOHNSON, R. A. & WICHERN, D. W. (1982)** *Applied Multivariate Statistical Analysis*. 2ª ed. Prentice Hall International. USA.
- KRZANOWSKI, W. J. (1984)** Principal Camponent Analysis in the Presence of Group Structure. *Appl. Statist.* 33, n° 2, pp 164-168.
- \_\_\_\_\_ (1979) Between-Groups composison of Components. *Journal of the Ammerican Statistical Association (JASA)* 74, n° 367, pp 703-707.

## BIBLIOGRAFIA

---

**KRZANOWSKI, W. J. and MARRIOTT, F. H. C. (1994)** *Multivariate Analysis. Part 1: Distributions, Ordination and Inference*. Edward Arnold. London.

\_\_\_\_\_ (1994) *Multivariate Analysis. Part 2: Classification, Covariance Structures and Repeated Measurements*. Edward Arnold. London.

**LANGRAND, C. (1996)** *Análises Fatoriais: Métodos de Classificação*. Seminário de Métodos Estatísticos Aplicados às Ciências Humanas. UFSCar/ULB. São Carlos.

**LEBART, L.; MORINEAU, A. and WARWICK, K. (1984)** *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices*. John Wiley & Sons. New York.

**LEÓN, M. A. V. (1993)** *Discriminação com Mistura de Variável Contínua e Categórica*. Campinas. IMECC/UNICAMP. Dissertação de Mestrado.

**LIMA FILHO, E. C. (1987)** *Planejamento de Observações Adicionais em Modelos Lineares: A Decomposição em Valores Singulares da Matriz de Planejamento e sua caracterização*. Campinas. IMECC/UNICAMP. Relatório Técnico N° 24/87.

**MAQUESA SOSA, B. R. (1989)** *Estudo de Tabelas de Contingência através de Análise de Correspondência*. Campinas. IMECC/UNICAMP. Dissertação de Mestrado.

**MARRIOTT, F. H. C. (1974)** *The Interpretation of Multiple Observations*. Academic Press. London.

**MARDIA, K.V; KENT, J.T. & BIBBY, J.M. (1979)** *Multivariate Analysis*. Academic Press. London.

## BIBLIOGRAFIA

---

- MOISÉS ORTEGA, E. M. (1997).** *Modelos de Preferência via Escalonamento Multidimensional*. Campinas. IMECC/UNICAMP. Tese de Mestrado.
- NISHISATO, S. (1993)** On Quantifying Different types of Categorical Data. *Psychometrika* **58** nº 4 pp. 617-629.
- OLIVEIRA, M. C. F. A. de,(1992)** Condição Feminina e Alternativa de Organização Doméstica: As Mulheres sem Compamheiro em São Paulo. *Revista Estudos Demográficos y Urbanos*. Cidade do México, nº 20-21, nov-dez/92.
- PEARSON, K. (1901)** On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, **2**, 559-572.
- PINHO, A. L. S. de, VIVACQUA, C. A., OLIVEIRA, M. C. F. A de, et all (1997)** *Projeto: Perfil da Organização Doméstica de Pessoas sem Companheiro em São Paulo*. Campinas. NEPO/UNICAMP. Relatório Técnico.
- PRADO, B. B. A. (1996)** *Introdução à Análise de Dados*. Seminário de Métodos Estatísticos Aplicados às Ciências Humanas. UFSCar/ULB. São Carlos.
- SAS Institute Inc. (1989)** SAS/STAT Users Guide. Version 6, Fourth Edition, Volume 1, Cary, NC; SAS Institute Inc., 943pp.
- SEARLE, S. R. (1982)** *Matrix Algebra Usefull for Statistics*. John Wiley & Sons. New York.
- SOUZA, A. M. R. de (1982)** *Análise de Correspondência*. . São Paulo. IME/USP. Dissertação de Mestrado.

BIBLIOGRAFIA

---

- TAKEUCHI, K., YANAI, H. and MUKHERJEE, B. N. (1982)** *The Foundations of Multivariate Analysis. A Unified Approach by Means of Projection Onto Linear Subspaces.* Wiley. New Delhi.
- TAKANE, Y. (1987)** Analysis of Contingency Tables by Ideal Point Discriminant Analysis. *Pschometrika* 52, nº 4, pp 493-513.
- TNENHAUS, M. and YOUNG, F. W. (1985)** *An Analysis and Synthesis of Multiple Correspondence Analysis, Optmal Scaling, Dual Scaling, Homogeneity Analysis and Other methods for quantifying Categorical Multivariate Data.* *Psychometryka* 50, nº 1. pp 91-119.
- VAN RIJCKEVORSEL (1986)** About Horseshoes in Multiple Correspondence Analysis. IN Gaul W. & Schader M. *Classification as a tool of research.* Elsevier Scence Publishers B. V. (North-Holland)
- VERDINELLI, M. A. (1980)** *Análise Inercial em Ecologia.* São Paulo. IO/USP. Tese de Doutorado.
- VILLALOBOS AGUAIO, M.T. (1993)** *Análise de Correspondência e Modelos Log-Lineares: um enfoque Integrado para Análise e Exploratória de Dados Categóricos.* Campinas. IMECC/UNICAMP. Dissertação de Mestrado.