

Universidade Estadual de Campinas
Instituto de Matemática Estatística e Ciência da Computação
Departamento de Estatística

**Análise de Correspondência
e Modelos Log-Lineares:
um Enfoque Integrado para a Análise
Exploratória de Dados Categóricos**

Maria Teresa Villalobos Aguayo

**Orientadora: Profa. Dra.
Regina Célia Carvalho Pinto Moran**

Campinas – São Paulo

Julho de 1993

Análise de Correspondência e Modelos Log-Lineares: um Enfoque Integrado para a Análise Exploratória de Dados Categóricos

Este exemplar corresponde à redação final da tese devidamente corrigida e defendida pela Sra. Maria Teresa Villalobos Aguayo e aprovada pela Comissão Julgadora.

Campinas, 10 de Setembro de 1993



Profa. Dra. Regina Célia Carvalho Pinto Moran
Orientadora

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação da Universidade Estadual de Campinas, como requisito parcial, para obtenção do Título de Mestre em Estatística.

Dedicação

A meu Márcio querido,

aos meus pais José Antonio e Teresa,

e aos meus sogros José Carlos e Maria do Carmo.

Agradecimentos

Quero agradecer a todos aqueles que colaboraram para a realização desta, em especial:

A Regina Moran, pela orientação, incentivo, apoio e amizade.

Ao Márcio, pelo amor e carinho e pela paciência.

A minha família, José Antonio, Teresa, Javier, Eduardo, Jose, Jorge e Liliana, pela presença constante, mesmo na distância.

A José Carlos, Maria do Carmo, tia Dinha, Maria José, Toninho, Maria Cecilia e Ana Lúcia, por serem minha família brasileira.

Aos funcionários, professores e companheiros do IMECC, pela cooperação e amizade.

A Matilde Scaramucci, pela colaboração cedendo um conjunto de dados.

Aos professores Jandyra M. G. Fachel, Aida C. G. Verdugo Lazo e Petrônio Pulino, pela participação na banca examinadora.

As instituições: CNPq e FUNCAMP, pelo apoio financeiro.

Resumo

Na literatura, a *Análise de Correspondência* (AC), é proposta como uma técnica para a análise exploratória de dados, mas pouco se fala de como ela pode integrar-se a um modelo log-linear, e servir como ferramenta para a interpretação das interações ou dos resíduos. Neste trabalho, a AC e uma das suas generalizações, *Análise de Correspondência Generalizada* (ACG), são desenvolvidas amplamente. A AC e ACG, são apresentadas como uma ferramenta que pode ser usada para a análise de resíduos de um modelo log-linear. Isto é possível reconhecendo as relações entre a AC e modelos log-lineares. Mais ainda, é mostrado como estas relações podem ser usadas para chegar a uma abordagem combinada para a análise de tabelas de contingência usando ambas as análises. No caso de tabelas multi-entradas, estas são recodificadas como tabelas de duas entradas, e logo analisadas com AC; neste caso, a AC pode ser interpretada como a decomposição de resíduos de um modelo log-linear específico de independência. A AC, também pode ser usada para decompor resíduos do modelo de quase-independência. A ACG, pode ser usada para decompor resíduos de modelos log-lineares menos restritos que o modelo de independência, como o modelo de simetria.

Sumário

1. Introdução	1
2. Análise de Correspondência	2
2.1 Caracterização da Análise de Correspondência	3
2.1.1 Introdução	3
2.1.2 Os Dados	5
2.1.3 Considerações da Análise	7
2.2 Configuração e Geometria dos Pontos	8
2.2.1 Formação dos Pontos no Espaço	9
2.2.2 As Massas	10
2.2.3 A Distância	10
2.2.4 Os Centróides	12
2.2.5 A Inércia	13
2.2.6 Resumindo a Notação Matricial	14
2.2.7 A Geometria	15
2.3 Critério do Bom Ajuste e Obtenção dos Eixos	19
2.3.1 A Análise em $\mathbb{R}^J$	19
2.3.2 A Análise em $\mathbb{R}^I$	24
2.3.3 Relações entre os Espaços	26
2.3.4 Reconstrução da Tabela de Contingência	31
2.4 Interpretação dos Resultados	33
2.4.1 Interpretação dos Pontos	33
2.4.2 Auxílios à Interpretação	40
2.4.3 Interpretação dos Eixos	43

2.4.4 Os Elementos Suplementares	44
2.5 Análise de Correspondência Generalizada	46
2.5.1 Os Dados	46
2.5.2 A Análise	49
2.5.3 Casos Particulares	51
 3. Relacionando a Análise de Correspondência com Modelos Log-Lineares	 52
3.1 Introdução	52
3.2 Análise Log-linear	53
3.3 Relações entre a Análise de Correspondência e alguns Modelos Log-lineares	55
3.4 Análise de Correspondência como uma Técnica para Decompor Resíduos de alguns Modelos de Independência Log-linear para Tabelas Multi-Entrada	58
3.5 Análise de Correspondência para Modelos de Quase-Independência	61
3.6 Análise de Correspondência para Decompor Resíduos do Modelo de Simetria	64
3.7 Análise de Correspondência para Decompor Resíduos do Modelo de Quase-Simetria	66
3.8 Observações Gerais	68
 4. Alguns Exemplos de Aplicação	 69
4.1 Aplicação da AC a um Exemplo Simples de Linguística	70
4.2 Aplicação da AC como um suplemento da aplicação da ALL a um Exemplo de Pena de Morte	77
4.2.1 Análise da Tabela	78
4.2.2 Análises das Tabelas Marginais	85

4.3 Aplicação da AC e ACG a um exemplo de Etologia	90
4.3.1 Analisando a Tabela com AC para a Decomposição dos resíduos do modelo de Independência	91
4.3.1 Analisando a Tabela com ACG para a Decomposição dos resíduos do modelo de Quase-Independência	96
4.3.1 Analisando a Tabela com ACG para a Decomposição dos resíduos do modelo de Simetria	101
5. Discussão Geral e Conclusões	105
Apêndices	107
A. A Distância Qui-Quadrado	108
B. Maximização de uma Forma Quadrática sob uma Restrição Quadrática	112
C. Decomposição em Valores Singulares	116
Referências Bibliográficas	120
Anexos	125
I. Exemplo da Lingüística	126
II. Exemplo da Pena de Morte	128
III. Exemplo da Etologia	135

1. Introdução

O presente trabalho, tem como objetivo a análise exploratória de dados categóricos, combinando as técnicas da *Análise de Correspondência* (AC) e *Análise Log-Linear* (ALL). Veremos a AC como uma técnica que pode ser usada para examinar os resíduos de modelos log-lineares. O uso de uma generalização da AC tem um rol importante nesta abordagem. Será mostrado que AC e ALL estão intimamente relacionados.

No segundo capítulo, será apresentado o desenvolvimento completo e profundo da AC, inclusive, uma generalização dela, a chamada *Análise de Correspondência Generalizada* (ACG).

No terceiro capítulo, será apresentada uma breve referência a ALL, e alguns resultados indicando as relações entre AC e ALL. Estas relações são usadas para entender a decomposição de resíduos de modelos log-lineares de independência, de quase-independência, simetria e quase-simetria e outros modelos mais gerais para tabelas de contingência.

No quarto capítulo, apresentaremos a aplicação da AC e ACG em exemplos reais, um simple e dois já trabalhados anteriormente por outros estatísticos.

Finalmente, no último capítulo, serão apresentadas uma discussão geral e as conclusões.

Em forma adicional, este trabalho compreende o Apêndice onde se trata:

- a) a distância qui-quadrado com suas propriedades e algumas observações,
- b) a maximização de uma forma quadrática sob uma restrição quadrática e
- c) alguns aspectos da decomposição de uma matriz.

2. Análise de Correspondência

"...A graphical description is more easily assimilated and interpreted than a numerical one and can assist all three functions 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 word which grows more complex and varied each day..."

Greenacre (1984)

2.1 Caracterização da Análise de Correspondência

2.1.1 Introdução

A Análise de Correspondência (a qual nos referiremos como AC), tem como objetivo, o estudo de *Tabelas de Freqüências Cruzadas* chamadas também de *Tabelas de Contingência*. A AC é uma técnica da análise exploratória, com a qual podemos gerar hipóteses a partir da análise da tabela, pois a AC fornece uma idéia da associação entre as variáveis categóricas envolvidas na tabela de contingência e a relação entre linhas e colunas, o que é especialmente útil quando as tabelas são grandes. Logo, com experimentos confirmatórios posteriores, poderemos testá-las empregando as metodologias usuais da estatística.

Um conjunto de variáveis categóricas com muitas categorias, geralmente, contém várias inter-relações que são impossíveis de interpretar a primeira vista. A abordagem da AC está enfatizada em representações geométricas, revelando a estrutura dos dados de forma ótima (com o que não há necessidade de assumir modelos nem distribuições fundamentais). Isto é possível, graças a uma adaptação da Análise de Componentes Principais, onde a representação é feita de tal forma que as distâncias entre pontos linha e ou entre pontos coluna no espaço euclidiano são iguais as chamadas distâncias χ^2 . Assim, a AC é considerada como um algoritmo de redução de dados que fornece imagens simplificadas da realidade multidimensional, e a busca da melhor representação simultânea que indica a dependência dos dois conjuntos de categorias, compreendidos pelas linhas e colunas da tabela de contingência. As regras de interpretação, estarão ligadas aos princípios geométricos das transformações efetuadas sobre os dados e nos permitirão remontar às estruturas reais a partir de estruturas representadas.

Para fixar idéias, o método poderia ser comparado ao uso de um aparelho radiográfico, que fornece imagens a partir de uma realidade não observável. O uso de tal aparelho para um diagnóstico, supõe uma certa preparação, por um lado, do doente que deve, por exemplo, absorver um produto, e pelo outro, de buscar a posição e direção mais convenientes para que, ao tirar a radiografia desde a posição e na direção, possamos captar as formas dos órgãos ou ossos

de interesse distribuídos no interior do corpo. A interpretação dos resultados está ligada ao conhecimento dos princípios de funcionamento do aparelho e a opacidade aos raios X que depende da densidade, do volume e da composição química dos órgãos, inclusive quando o sujeito absorve o produto.

Podemos distinguir AC *Simple*s (AC de tabelas de contingência), AC *Generalizada* (ACG) (uma generalização da AC), e AC *Múltipla* (ACM)(AC de matrizes indicadoras e/ou matrizes de Burt). Atualmente, AC simples é conhecida sob diferentes nomes, tais como *Ponderação Recíproca*, *Dual (ou Optimal) Scaling*, *Análise de Correlação Canônica*, e *Regressão Linear Simultânea*; estes diferentes nomes são devido aos diferentes contextos ou abordagens sob os quais foram criados. Greenacre (1984), discute as diferenças entre AC simples e estes métodos, mas mostra a equivalência destes com AC (veja também Nishisato (1980) capítulo 3). A AC, trabalha basicamente com a diferença entre as observações e a esperança sob hipótese de independência; no caso da ACG, a generalização é voltada para a diferença entre as observações e a esperança sob algum modelo em geral. Por outro lado, a ACM é também conhecida com outros nomes como *Análise de Homogeneidade*, *Dual Scaling* ou *Optimal Scaling* (veja Tenenhaus e Young (1985)).

A História da Análise de Correspondência

Aparentemente, a história da AC começa em 1935, quando H.O.Hartley, sob o seu nome alemão original Hirschfeld, publicou *A Connection Between Correlation and Contingency*, onde ele propôs uma solução algébrica para a correlação entre linhas e colunas de uma tabela de contingência. Na mesma época, idéias similares, mas com uma abordagem diferente, foram sugeridas por Richardson e Kuder (1933) e depois de forma independente por Horst (1935), que criou o termo *Method of Reciprocal Averaging* para aplicação na psicometria. Posteriormente, Fisher (1940), no seu artigo *The Precision of Discriminant Functions*, desenvolveu a teoria fundamental e propôs uma análise discriminante para tabelas de contingência. Por outro lado, Guttman (1941), desenvolveu independentemente um método para construir escalas para dados categóricos que, partindo de um contexto diferente, resultou ser um método equivalente ao da AC. Desde então, foi redescoberta em diferentes circunstâncias e também praticamente esquecida, até a popularização do método que teve lugar em meados da década de 1960, graças a um grupo de franceses liderados por Jean-Paul Benzécri. O grupo, desenvolveu o método como uma

técnica gráfica multidimensional, cujas implicações geométricas facilitaram a interpretação. E deu ao método, o nome de "*Analyse Factorielle des Correspondances*", para denotar com o termo "*correspondance*" a um "*sistema de associações*" entre os elementos de dois conjuntos de categorias; especificamente, quando Benzécri analisou uma tabela de consoantes contra vocais finais, da lingüística chinesa (Greenacre, 1984). Pouco depois o nome foi simplificado a "*Analyse des Correspondances*". A partir daí, temos uma grande quantidade de literatura no assunto por: Lebart e Fenelon (1971), Benzécri (1973), Hill (1974), Nishisato (1980), Gifi (1981), Greenacre (1984), Lebart, Morineau e Warwick (1984), Everitt (1992), etc... Em português, Verdinelli (1980), Souza (1982), Maquera Sosa (1989), Souza (1990), Cabral (1991) e Callegari-Jaques (1991). Desde 1976, existe um jornal dedicado quase exclusivamente às aplicações da AC, chamado *Les Cahiers de l'Analyse des Données*.

Este capítulo, se ocupa amplamente do desenvolvimento da AC, e está baseado principalmente, nos trabalhos de Grande e Abascal (1989), Greenacre (1984), Lebart, Morineau e Warwick (1984), Greenacre e Hastie (1987) e Goodman (1986, 1991). Neste capítulo, veremos como são apresentados os dados, e sob quais considerações analisaremos estes. Na segunda seção, a configuração e geometria dos pontos. Na terceira, todo o referente a análise em si: O critério do bom ajuste e obtenção dos eixos, que está sub-dividido em: a análise em $\mathbb{R}^I$, a análise em $\mathbb{R}^I$, as relações entre os espaços, e reconstrução da tabela de contingência. Na quarta, a interpretação dos resultados. Finalmente, a última seção, trata da Análise de Correspondência Generalizada.

2.1.2 Os Dados

Geralmente, os dados são representados sob a forma de uma tabela de dupla entrada, na qual figuram:

- Em linhas, I categorias de uma variável categórica:

$$Cat'_1, Cat'_2, \dots, Cat'_I$$

- Em colunas, J categorias de uma outra variável categórica:

$$\text{Cat}^*_1, \text{Cat}^*_2, \dots, \text{Cat}^*_J$$

- Designaremos por n_{ij} ($1 \leq i \leq I, 1 \leq j \leq J$) ao número de casos observados que pertencem a intersecção da i -ésima categoria da primeira variável e da j -ésima categoria da segunda variável.

	Cat^*_1	Cat^*_2	$\dots$	Cat^*_j	$\dots$	Cat^*_J
Cat'_1				.		
Cat'_2				.		
$\dots$				.		
Cat'_i	.	.	.	n_{ij}		
$\dots$						
Cat'_I						

A tabela de contingência é a matriz:

$$\mathbf{N} = [n_{ij}]_{I \times J} \quad (1)$$

de ordem $I \times J$, onde, $n_{ij} \geq 0$, para todo $i = 1, \dots, I$, e $j = 1, \dots, J$.

O valor n_{ij} representa a frequência conjunta das categorias i e j . Com o propósito de simplificar a linguagem, usaremos de agora em diante o termo "*linha*" quando nos referirmos a alguma categoria representada numa linha de $\mathbf{N}$, e usaremos o termo "*coluna*" quando nos referirmos a alguma categoria representada numa coluna de $\mathbf{N}$. Definiremos:

$$\begin{aligned} n_{i.} &= \sum_j n_{ij} && , \text{ o total da linha } i \\ n_{.j} &= \sum_i n_{ij} && , \text{ o total da coluna } j \\ n &= \sum_{ij} n_{ij} && , \text{ o total absoluto.} \end{aligned} \quad (2)$$

Definiremos a *Matriz de Freqüências Relativas* como:

$$\mathbf{P} = 1/n \mathbf{N}, \quad (3)$$

esta é chamada por alguns autores de *Matriz de Correspondência*. Devido ao fato desta matriz estar constituída unicamente por entradas não negativas, pode ser considerada como a distribuição de uma unidade de massa entre as caselas. Esta é a matriz sobre a qual realizaremos a nossa análise. Temos que:

$$\begin{aligned} p_{i.} &= \sum_j p_{ij} = n_{i.}/n && , \text{ a freqüência relativa da linha } i \\ p_{.j} &= \sum_i p_{ij} = n_{.j}/n && , \text{ a freqüência relativa da coluna } j, \text{ e} \\ p_{..} &= \sum_{ij} p_{ij} = 1. \end{aligned} \quad (4)$$

2.1.3 Considerações da Análise

Pensemos nas considerações que deveremos fazer para nossa análise. Por um lado, a análise não deverá fazer diferença quando transpusermos a matriz de contingência. Por outro lado, não interessa ver as diferenças absolutas que existem entre duas linhas (ou duas colunas) e sim os padrões de distribuição que adotam diante das colunas (linhas), e consideraremos que duas linhas (ou duas colunas) são semelhantes se apresentam a mesma distribuição condicional. Portanto, o método buscado deverá:

- Ser simétrico com relação às linhas e colunas.
- Trabalhar com as distribuições de freqüências condicionais.

Além disso, se queremos investigar como se relacionam linhas e colunas, observemos que:

- Se numa linha i encontrássemos que p_{ij} é grande com relação aos restantes da linha, poderíamos pensar que a linha i está caracterizada pela coluna j . Mas, isto só estaria correto, se o valor p_{ij} fosse maior do que o esperado sob hipótese de independência (i.e. se $p_{ij} > p_{i.}p_{.j}$, ou visto de outro modo com relação a distribuição condicional: $p_{ij}/p_{i.} > p_{.j}$). Em outro caso,

não poderemos obter uma conclusão sobre a linha i em relação à coluna j ; sómente que este elemento toma valores altos em geral.

- De forma análoga e equivalente, se numa coluna j encontrássemos que p_{ij} é maior que os restantes da coluna, poderíamos pensar que a coluna j está caracterizada pela linha i . Mas, isto só estaria correto, se o valor p_{ij} fosse maior do que o esperado sob hipótese de independência (i.e. se $p_{ij} > p_{i.}p_{.j}$, ou visto de outro modo com relação a distribuição condicional: $p_{ij}/p_{.j} > p_{i.}$). Em outro caso, não poderemos obter uma conclusão sobre a coluna j em relação à linha i ; somente que este elemento toma valores altos em geral.

Portanto, quando examinarmos as linhas ou colunas da tabela, não somente deveremos observar suas distribuições condicionais, mas também elas com relação as distribuições marginais.

2.2 Configuração e Geometria dos Pontos

Antes de começar a trabalhar com os dados, veremos algumas definições preliminares.

Definiremos os vetores $\mathbf{r} \in \mathbb{R}^I$ e $\mathbf{c} \in \mathbb{R}^J$ como sendo os vetores de frequências relativas linha e coluna respectivamente. Assim:

$$\begin{aligned}\mathbf{r} &= (p_{1.}, p_{2.}, \dots, p_{I.})' = \mathbf{P} \mathbf{1}_J \\ \mathbf{c} &= (p_{.1}, p_{.2}, \dots, p_{.J})' = \mathbf{P}' \mathbf{1}_I,\end{aligned}\tag{5}$$

onde $\mathbf{1}_J$ e $\mathbf{1}_I$ são vetores cujas coordenadas são todas uns, de dimensões J e I respectivamente. Os vetores $\mathbf{r}$ e $\mathbf{c}$ podem ser vistos como as distribuições marginais linha e coluna respectivamente.

Definiremos também as matrizes diagonais de frequências relativas $\mathbf{D}_r$ e $\mathbf{D}_c$ de dimensões I e J respectivamente:

$$\begin{aligned}\mathbf{D}_r &= \text{diag}\{\mathbf{r}\} \\ \mathbf{D}_c &= \text{diag}\{\mathbf{c}\}.\end{aligned}\tag{6}$$

2.2.1 Formação dos Pontos no Espaço

Seguindo as considerações feitas na seção 2.1.3, o interesse estará em representar as categorias por vetores, cujas coordenadas sejam as distribuições condicionais.

- Em $\mathbb{R}^J$ formaremos uma nuvem de I pontos, que representarão as distribuições condicionais por linha e que chamaremos de *Perfis Linha*. Definiremos a matriz que contém nas linhas os perfis linha como:

$$\mathbf{R} = \mathbf{D}_r^{-1} \mathbf{P},\tag{7}$$

isto é: $\mathbf{R} = [\mathbf{r}_1 \mid \mathbf{r}_2 \mid \dots \mid \mathbf{r}_I]'$, onde para todo i ($i = 1, \dots, I$):

$$\mathbf{r}_i = (p_{i1}/p_{i.}, p_{i2}/p_{i.}, \dots, p_{iJ}/p_{i.})'.$$

- Em $\mathbb{R}^I$ formaremos uma nuvem de J pontos, que representarão as distribuições condicionais por coluna e que chamaremos de *Perfis Coluna*. Definiremos a matriz que contém nas linhas os perfis coluna como:

$$\mathbf{C} = \mathbf{D}_c^{-1} \mathbf{P}',\tag{8}$$

isto é: $\mathbf{C} = [\mathbf{c}_1 \mid \mathbf{c}_2 \mid \dots \mid \mathbf{c}_J]'$, onde para todo j ($j = 1, \dots, J$):

$$\mathbf{c}_j = (p_{1j}/p_{.j}, p_{2j}/p_{.j}, \dots, p_{Ij}/p_{.j})'$$

Vemos que, as transformações realizadas a partir da matriz $\mathbf{P}$, são do mesmo tipo nos dois espaços $\mathbb{R}^I$ e $\mathbb{R}^J$. Como pode ser observado, a matriz dos dados em $\mathbb{R}^I$ não é a transposta da matriz dos dados em $\mathbb{R}^J$, como é o que acontece na Análise de Componentes Principais (veja Lebart, Morineau e Warwick, 1984, p. 10-14). Isto, nos conduzirá a realizar duas análises diferentes, uma em cada espaço. Felizmente, encontraremos algumas relações que permitirão a redução dos cálculos, e também facilitarão a interpretação.

2.2.2 As Massas

Embora na AC trabalhemos com perfis, não devemos esquecer as diferenças entre os totais de cada linha (ou de cada coluna), por isso atribuiremos a cada perfil uma massa proporcional à sua frequência, e diremos que:

- Em $\mathbb{R}^I$ cada ponto i ($i = 1, \dots, I$) está afetado por uma massa $p_{i.}$, assim de (5), temos que o vetor das massas das linhas é $\mathbf{r}$.
- Em $\mathbb{R}^J$ cada ponto j ($j = 1, \dots, J$) está afetado por uma massa $p_{.j}$, assim de (5), o vetor das massas das colunas é $\mathbf{c}$.

Notemos que em cada espaço, a soma das massas é 1, i.e.: $\mathbf{1}_I' \mathbf{r} = \mathbf{1}_J' \mathbf{c} = 1$.

2.2.3 A Distância

O fato de trabalhar com distribuições condicionais, nos conduz a utilizar uma distância distinta da euclidiana usual. A AC utiliza uma distância apropriada para distribuições, a distância *Qui-quadrado* que denotaremos por χ^2 .

- A distância χ^2 entre dois perfis linha i e l ($i, l = 1, \dots, I$):

$$\begin{aligned} d^2_{\{\mathbf{R}\}}(i, l) &= \sum_j 1/p_{.j} (p_{ij}/p_{i.} - p_{lj}/p_{l.})^2 \\ &= (\mathbf{r}_i - \mathbf{r}_l)' \mathbf{D}_c^{-1} (\mathbf{r}_i - \mathbf{r}_l) . \end{aligned} \tag{9}$$

- A distância χ^2 entre dois perfis coluna j e k ($j, k = 1, \dots, J$):

$$\begin{aligned} d^2_{[c]}(j, k) &= \sum_i 1/p_i \cdot (p_{ij}/p_{.j} - p_{ik}/p_{.k})^2 \\ &= (\mathbf{c}_j - \mathbf{c}_k)' \mathbf{D}_r^{-1} (\mathbf{c}_j - \mathbf{c}_k) . \end{aligned} \quad (10)$$

esta distância é uma distância euclidiana, mas difere da euclidiana usual unicamente na ponderação, sendo $\mathbf{D}_c^{-1}$ e $\mathbf{D}_r^{-1}$ nos espaços $\mathbb{R}^I$ e $\mathbb{R}^J$ respectivamente.

A fórmula em (9), mostra que $d^2_{[R]}(i, l)$ é uma medida para a diferença entre os perfis i e l . Quando i e l têm o mesmo perfil, então $d^2_{[R]}(i, l) = 0$.

Ao examinar a fórmula da distância entre linhas, poderemos dizer: se a diferença $(p_{ij}/p_i - p_{lj}/p_l)$ for pequena com relação à massa da coluna j ($p_{.j}$), então, por meio da ponderação por $(1/p_{.j})$, estaremos evitando que esta diferença tenha muita influência na distância; por outro lado, se tal diferença fosse grande com relação à massa, então, com a ponderação, influenciará de forma proporcional na distância. Poderemos raciocinar de forma análoga quando observarmos a distância entre colunas.

Utilizando a distância χ^2 , estaremos dando um efeito estabilizador sobre os dados, visto que cumpre a propriedade denominada *Equivalência Distribucional*, que se expressa da seguinte maneira:

- Se duas linhas têm o mesmo perfil, elas podem ser substituídas por uma só, a qual tem o mesmo perfil, mas com uma massa igual a soma das massas, sem que isto altere as distâncias entre os perfis coluna.
- Se duas colunas têm o mesmo perfil, elas podem ser substituídas por uma só, a qual tem o mesmo perfil, mas com uma massa igual a soma das massas, sem que isto altere as distâncias entre os perfis linha.

Assim, se duas categorias com perfis iguais são reunidas em uma só, os resultados em geral não serão modificados. Se existirem dois perfis iguais, eles serão vistos como um único ponto, correspondente às duas categorias unidas; por um ponto de vista técnico, não ganharíamos nada no estudo das duas variáveis envolvidas, se sub-dividimos uma categoria em categorias homogêneas. Por outro lado, se categorias com perfis parecidos fossem

juntadas, a representação dos pontos mudaria pouco. A demonstração da propriedade de equivalência distribucional, está no apêndice A.

A propriedade de equivalência distribucional não é a única base para reconhecer a distância χ^2 como apropriada para distribuições. A distância χ^2 está intimamente relacionada à estatística χ^2 pois, a soma ponderada das distâncias χ^2 dos pontos ao seu centróide (que veremos em seguida) é igual a estatística χ^2 dividida por n (veja seções 2.2.4 e 2.2.5). Assim, a configuração dos pontos com estas distâncias representa uma decomposição da estatística χ^2 .

A seguir, apresentaremos dois conceitos elementares da configuração dos pontos.

2.2.4 Os Centróides

Ao calcular os centróides não deveremos esquecer as massas correspondentes a cada ponto, e definiremos;

- O centróide dos perfis linha por:

$$\text{Centróide}_{\{\mathbf{R}\}} = \sum_i p_i \cdot \mathbf{R}_i = \mathbf{R}' \mathbf{r} , \quad (11)$$

substituindo $\mathbf{R}$ por sua expressão em (7), temos:

$$\text{Centróide}_{\{\mathbf{R}\}} = \mathbf{c} .$$

- Análogamente, o centróide dos perfis coluna está definido por:

$$\text{Centróide}_{\{\mathbf{C}\}} = \sum_j p_{.j} \mathbf{C}_j = \mathbf{C}' \mathbf{c} , \quad (12)$$

substituindo $\mathbf{C}$ por sua expressão em (8), temos:

$$\text{Centróide}_{\{\mathbf{C}\}} = \mathbf{r} .$$

2.2.5 A Inércia

A Inércia é a quantificação da variação total de cada nuvem de pontos.

- A inércia dos perfis linha se define como a soma ponderada pela massa, das distâncias X^2 dos pontos ao centróide, e a denotaremos por $In_{[R]}$, i.e.:

$$In_{[R]} = \sum_i p_i \cdot d^2_{[R]}(i, c). \quad (13)$$

- Análogamente, a inércia dos perfis coluna se define como a soma ponderada pela massa, das distâncias X^2 dos pontos ao centróide, e a denotaremos por $In_{[C]}$, i.e.:

$$In_{[C]} = \sum_j p_{.j} \cdot d^2_{[C]}(j, r). \quad (14)$$

Claramente, esta é uma *Medida de Variação Total* ou *Variância Total*, na que representamos os perfis como se fossem indivíduos, e as categorias como se fossem variáveis, através das quais estes perfis se distribuem. O termo *Inércia* é análogo à definição matemática de *Momento de Inércia*, que é a integral de massa multiplicada pela distância quadrada ao centróide. Aparentemente, este nome foi adotado pelos franceses, devido à ênfase na geometria e na física, pois lembra a um sistema de massas pontuais. Na literatura inglesa, a inércia é conhecida como o *Coeficiente Quadrado Médio de Contingência de Pearson* calculado sobre **N**.

Das expressões (13) e (14) obtemos que:

$$\begin{aligned} In_{[R]} &= \sum_i p_i \cdot [(\mathbf{R}_i - \mathbf{c})' \mathbf{D}_c^{-1} (\mathbf{R}_i - \mathbf{c})] \\ &= \text{tr} [\mathbf{D}_r (\mathbf{R} - \mathbf{1c}') \mathbf{D}_c^{-1} (\mathbf{R} - \mathbf{1c}')'] \\ &= \text{tr} [(\mathbf{P} - \mathbf{rc}') \mathbf{D}_c^{-1} (\mathbf{P} - \mathbf{rc}')' \mathbf{D}_r^{-1}] \end{aligned} \quad (15)$$

$$\begin{aligned}
e \quad In_{[c]} &= \sum_j p_{.j} [(\mathbf{c}_j - \mathbf{r})' \mathbf{D}_r^{-1} (\mathbf{c}_j - \mathbf{r})] \\
&= \text{tr} [\mathbf{D}_c (\mathbf{C} - \mathbf{1} \mathbf{r}') \mathbf{D}_r^{-1} (\mathbf{C} - \mathbf{1} \mathbf{r}')'] \quad (16) \\
&= \text{tr} [(\mathbf{P}' - \mathbf{c} \mathbf{r}') \mathbf{D}_r^{-1} (\mathbf{P}' - \mathbf{c} \mathbf{r}')' \mathbf{D}_c^{-1}] \\
&= \text{tr} [(\mathbf{P} - \mathbf{r} \mathbf{c}')' \mathbf{D}_r^{-1} (\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{D}_c^{-1}].
\end{aligned}$$

Com ambos os resultados, temos que a inércia das linhas é igual a inércia das colunas (veja Mardia, 1979, p.456, seção A.2.2b), e conseqüentemente a chamaremos simplesmente de *Inércia* e a denotaremos por In .

Também temos que In é igual a usual estatística χ^2 dividida por n , pois:

$$\begin{aligned}
In &= \text{tr} [(\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{D}_c^{-1} (\mathbf{P} - \mathbf{r} \mathbf{c}')' \mathbf{D}_r^{-1}] \\
&= 1/n \text{tr} [1/n n(\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{D}_c^{-1} n(\mathbf{P} - \mathbf{r} \mathbf{c}')' \mathbf{D}_r^{-1}] \quad (17) \\
&= 1/n \text{tr} [1/n (\mathbf{N} - \mathbf{E}) \mathbf{D}_c^{-1} (\mathbf{N} - \mathbf{E})' \mathbf{D}_r^{-1}] \\
&= 1/n \chi^2,
\end{aligned}$$

onde $\mathbf{E}$ é a Esperança de $\mathbf{N}$ sob hipótese de independência, i.e.: $\mathbf{E} = n \mathbf{r} \mathbf{c}'$.

Daqui temos que, com o uso da distância χ^2 , a AC "decompõe" a estatística χ^2 , de modo que, quando a estatística χ^2 indicar afastamento do modelo de independência, poderemos reconhecer por causa de quais categorias (linhas e colunas) isto acontece. O uso deste fato será melhor explicado na seção 2.4.2.

2.2.6 Resumindo a Notação Matricial

Matriz dos Dados	:	$\mathbf{N} \equiv [n_{ij}]_{I \times J}$	, onde $n_{ij} \geq 0$
Matriz de Correspondência	:	$\mathbf{P} \equiv 1/n \mathbf{N}$	, onde $n = \mathbf{1}' \mathbf{N} \mathbf{1}$
Freqüências Relativas	:	$\mathbf{r} \equiv \mathbf{P} \mathbf{1}$	$\mathbf{c} \equiv \mathbf{P}' \mathbf{1}$
Diagonais de Freq. Relativas	:	$\mathbf{D}_r \equiv \text{diag}\{\mathbf{r}\}$	$\mathbf{D}_c \equiv \text{diag}\{\mathbf{c}\}$
Perfis Linha e Coluna	:	$\mathbf{R} \equiv \mathbf{D}_r^{-1} \mathbf{P}$	$\mathbf{C} \equiv \mathbf{D}_c^{-1} \mathbf{P}'$

Os perfis linha e coluna definem duas nuvens de pontos:

Nuvem Perfis Linha	Nuvem Perfis Coluna
Pontos: I perfis linha $\in \mathbb{R}^J$ $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_I$ Massas: I elementos de $\mathbf{r}$ Métrica: Euclidiana Ponderada definida por $\mathbf{D}_{\mathbf{c}}^{-1}$	Pontos: J perfis linha $\in \mathbb{R}^I$ $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_J$ Massas: J elementos de $\mathbf{c}$ Métrica: Euclidiana Ponderada definida por $\mathbf{D}_{\mathbf{r}}^{-1}$

2.2.7 A Geometria

Nos limitaremos a examinar a geometria dos pontos em um dos espaços, neste caso em $\mathbb{R}^J$, já que as observações no outro espaço $\mathbb{R}^I$, podem ser deduzidas, de forma análoga, usando os elementos acima descritos para $\mathbb{R}^I$.

Os Perfis

Os perfis linha, estão representados nas linhas de $\mathbf{R} = \mathbf{D}_{\mathbf{r}}^{-1} \mathbf{P}$, então temos que:

- Como $p_{ij} \geq 0$, os pontos estão no primeiro quadrante de $\mathbb{R}^J$.
- Os perfis linha estão contidos no hiperplano $\pi_{\{\mathbf{R}\}}$ definido por:

$$\pi_{\{\mathbf{R}\}} = \{ \mathbf{x} \in \mathbb{R}^J : \mathbf{1}'(\mathbf{x} - \mathbf{c}) = 0 \}, \quad (18)$$

pois, para todo i , $i = 1, \dots, I$: $\mathbf{1}'(\mathbf{r}_i - \mathbf{c}) = 1 - 1 = 0$

Notemos que:

- O centróide $\mathbf{c}$ dos perfis encontra-se também em $\pi_{\{\mathbf{R}\}}$.
- $\pi_{\{\mathbf{R}\}}$ contém os pontos $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_J$, onde $\mathbf{e}_j$ é o j -ésimo vetor da base canônica de $\mathbb{R}^J$, pois para todo $j = 1, \dots, J$, temos: $\mathbf{1}'(\mathbf{e}_j - \mathbf{c}) = 1 - 1 = 0$.

- O vetor $\mathbf{1}$ é ortogonal ao plano $\pi_{\{\mathbf{R}\}}$, e todos os pontos linha de $\mathbf{R}$ têm uma projeção igual a $J^{-1/2}$ sobre o vetor. Com efeito, para todo $i = 1, \dots, I$, temos:

$$\text{proj} \{ \mathbf{R}_i \text{ sobre } \mathbf{1} \} = \mathbf{1}' \mathbf{R}_i / (\mathbf{1}' \mathbf{1})^{1/2} = 1/J^{1/2}.$$

Portanto, se cumpre que todos os perfis linha, estão no simplex de $\mathbb{R}^J$.

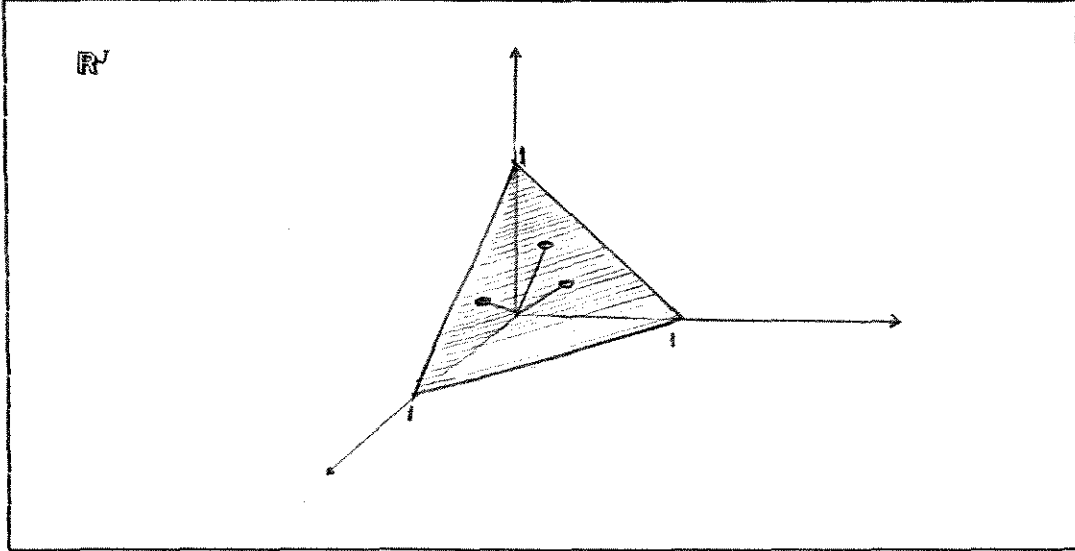


Figura 1: Os pontos de $\pi_{\{\mathbf{R}\}}$.

A Distância

Partindo da expressão (9), a distância entre as linhas i e l , pode ser expressada da seguinte maneira:

$$d_{\{\mathbf{R}\}}^2(i, l) = (\mathbf{D}_c^{-1/2} \mathbf{R}_i - \mathbf{D}_c^{-1/2} \mathbf{R}_l)' (\mathbf{D}_c^{-1/2} \mathbf{R}_i - \mathbf{D}_c^{-1/2} \mathbf{R}_l) \quad (19)$$

portanto, para visualizarmos a distância χ^2 entre os pontos $\mathbf{R}_i$ e $\mathbf{R}_l$, bastará visualizar a distância euclidiana dos pontos transformados $\mathbf{D}_c^{-1/2} \mathbf{R}_i$ e $\mathbf{D}_c^{-1/2} \mathbf{R}_l$.

Em consequência, realizar uma análise considerando a distância $\mathbf{x}^2$ é o mesmo que realizar uma análise com a distância euclidiana aplicada aos novos pontos linha da matriz:

$$\mathbf{RD}_c^{-\frac{1}{2}}, \quad (20)$$

mas sem deixar de considerar as massas correspondentes.

Os Perfis Transformados

Representaremos os perfis linha com os vetores linha da matriz transformada em (20). Partindo da expressão do hiperplano $\pi_{|\mathbf{R}|}$ em (18), obteremos equivalentemente a equação $[\mathbf{1}'\mathbf{D}_c^{-\frac{1}{2}}][\mathbf{D}_c^{-\frac{1}{2}}(\mathbf{x} - \mathbf{c})] = 0$, que resulta em $[\mathbf{D}_c^{\frac{1}{2}}\mathbf{1}]'[\mathbf{D}_c^{-\frac{1}{2}}\mathbf{x} - \mathbf{D}_c^{\frac{1}{2}}\mathbf{1}] = 0$. Portanto:

Os pontos linha da matriz transformada $\mathbf{RD}_c^{-\frac{1}{2}}$ estão no hiperplano ortogonal ao vetor $\mathbf{D}_c^{\frac{1}{2}}\mathbf{1}$ e que passa pelo ponto $\mathbf{D}_c^{\frac{1}{2}}\mathbf{1}$. i.e. o hiperplano definido por:

$$\pi_{|\mathbf{RD}_c^{-\frac{1}{2}}|} = \{ \mathbf{y} \in \mathbb{R}^J : (\mathbf{D}_c^{\frac{1}{2}}\mathbf{1})'(\mathbf{y} - \mathbf{D}_c^{\frac{1}{2}}\mathbf{1}) = 0 \}, \quad (21)$$

onde $\mathbf{y} = \mathbf{D}_c^{-\frac{1}{2}}\mathbf{x}$, com $\mathbf{x} \in \pi_{|\mathbf{R}|}$.

Notemos que:

- O centróide dos perfis transformados encontra-se também em $\pi_{|\mathbf{RD}_c^{-\frac{1}{2}}|}$, pois: $\text{Centróide}_{|\mathbf{RD}_c^{-\frac{1}{2}}|} = (\mathbf{RD}_c^{-\frac{1}{2}})'\mathbf{r} = \mathbf{D}_c^{\frac{1}{2}}\mathbf{1}$.
- O hiperplano $\pi_{|\mathbf{RD}_c^{-\frac{1}{2}}|}$ contém os pontos $(p_{.j}^{-\frac{1}{2}}\mathbf{e}_j)$, onde $\mathbf{e}_j$ é o j -ésimo vetor da base canônica de $\mathbb{R}^J$, pois para todo $j = 1, \dots, J$, temos que:
 $(\mathbf{D}_c^{\frac{1}{2}}\mathbf{1})'(p_{.j}^{-\frac{1}{2}}\mathbf{e}_j - \mathbf{D}_c^{\frac{1}{2}}\mathbf{1}) = 1 - 1 = 0$.
- O vetor unitário $(\mathbf{D}_c^{\frac{1}{2}}\mathbf{1})$ é ortogonal ao plano $\pi_{|\mathbf{RD}_c^{-\frac{1}{2}}|}$, e todos os pontos linha de $\mathbf{RD}_c^{-\frac{1}{2}}$ têm uma projeção igual a 1 sobre o vetor. Com efeito, para todo $i = 1, \dots, I$, temos:
 $\text{proj}\{\mathbf{D}_c^{-\frac{1}{2}}\mathbf{r}_i \text{ sobre } \mathbf{D}_c^{\frac{1}{2}}\mathbf{1}\} = (\mathbf{D}_c^{\frac{1}{2}}\mathbf{1})'(\mathbf{D}_c^{-\frac{1}{2}}\mathbf{r}_i) = 1$

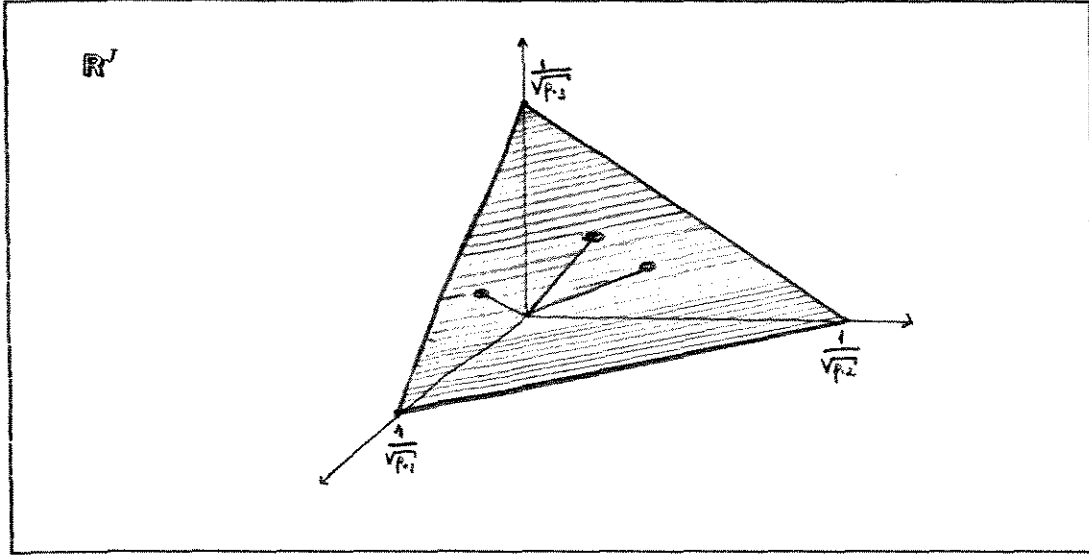


Figura 2: Os pontos de $\pi_{|RDc^{-1/2}|}$.

Depois de localizar os pontos no espaço, nosso objetivo será representar Graficamente os pontos $(D_c^{-1/2} \mathbf{R}_i)$ ($i = 1, \dots, I$), sobre um novo sistema de coordenadas, para que, ao observá-los, possamos ver a distância χ^2 , já que a distância euclidiana entre eles é a distância χ^2 entre os pontos $\mathbf{R}_i$. Portanto, nossos pontos serão as linhas de $RD_c^{-1/2}$.

Depois de centrarmos a nova nuvem, trasladando a origem ao seu centro de gravidade, as novas coordenadas dos pontos linha, estão dadas na matriz:

$$(\mathbf{R} - \mathbf{1} \mathbf{c}') D_c^{-1/2} \quad (22)$$

2.3 Critério do Bom Ajuste e Obtenção dos Eixos

Na AC as transformações feitas nos conduzem a realizar duas análises, uma em cada espaço.

2.3.1 A Análise em $\mathbb{R}^J$

O Problema Proposto

O objetivo é obter uma representação simplificada dos perfis linha, dando uma idéia das distâncias X^2 entre eles, portanto, tomando em conta as observações feitas na seção 2.2.7, representaremos os perfis pela matriz em (20): RD_c^{-1} .

O problema é resolvido identificando sub-espacos de $\mathbb{R}^J$ de menor dimensão que J , e tal que, a configuração dos I pontos tenha a menor deformação ao projetar-se nestes sub-espacos, de modo que as localizações verdadeiras dos I pontos possam ser cobertas satisfatoriamente desde suas posições neste sub-espaco. Começaremos por buscar um sub-espaco vetorial de uma dimensão, o qual represente o "melhor ajuste possível dos I perfis", no sentido em que a soma ponderada pela massa das distâncias quadradas dos pontos ao sub-espaco, seja mínima; logo, buscaremos um 2º sub-espaco unidimensional que seja ortogonal ao primeiro, mas com características similares, e assim por diante. O problema se reduz a buscar uma nova base ortonormal, tal que represente o melhor ajuste no sentido de quadrados mínimos. Logo, pelo método de Quadrados Mínimos, determinamos as coordenadas sobre o novo sistema. Assim a AC é considerada como uma adaptação do método de Componentes Principais (Pearson, 1901).

A Solução

Este método é equivalente a maximizar a inércia das projeções dos pontos sobre o sub-espaco com base $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ ($d = 1, \dots, J$), tal que:

- 1º) A inércia dos pontos projeção da nuvem sobre $\mathbf{v}_1$ seja máxima, e $\mathbf{v}_1' \mathbf{v}_1 = 1$.

.....

$d \circ$) A inércia dos pontos projeção da nuvem sobre $\mathbf{v}_d$ seja máxima, $\mathbf{v}_d' \mathbf{v}_d = 1$ e $\mathbf{v}_d' \mathbf{v}_{\alpha-1} = 0, \dots, \mathbf{v}_d' \mathbf{v}_1 = 0$.

Em tal sentido, está demonstrado que o sub-espaco d -dimensional ($d = 1, \dots, J$) que melhor se ajusta aos dados, contém necessariamente o centróide dos pontos (Greenacre, 1984, seções 2.5 - 2.6.3). Não incluiremos tal demonstração, pois mais adiante veremos que o fato de centrar os dados na AC, não é imprescindível. Uma vez visto que, o centróide encontra-se no sub-espaco optimal, e estaremos interessados em representar os pontos transformados da matriz em (22): $[(\mathbf{R} - \mathbf{1} \mathbf{c}') \mathbf{D}_c^{-1}]$.

Temos que, as projeções dos pontos transformados sobre um eixo de direção $\mathbf{v}$ de $\mathbb{R}^J$, unitário ($\mathbf{v}' \mathbf{v} = 1$), estão em:

$$\mathbf{f}(\mathbf{v}) = (\mathbf{R} - \mathbf{1} \mathbf{c}') \mathbf{D}_c^{-1} \mathbf{v}, \quad (23)$$

assim, a matriz das projeções sobre os eixos procurados $\mathbf{v}_1, \dots, \mathbf{v}_\alpha, \dots, \mathbf{v}_J$ é:

$$\mathbf{F} = (\mathbf{R} - \mathbf{1} \mathbf{c}') \mathbf{D}_c^{-1} \mathbf{V}, \quad (24)$$

onde: $\mathbf{V} = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \dots \mid \mathbf{v}_J]$, com $\mathbf{V}' \mathbf{V} = \mathbf{I}$ e

$$\mathbf{F} = [\mathbf{F}_{(1)} \mid \mathbf{F}_{(2)} \mid \dots \mid \mathbf{F}_{(J)}], \text{ com } \mathbf{F}_{(\alpha)} = \mathbf{f}(\mathbf{v}_\alpha).$$

A matriz $\mathbf{F}$ é chamada de *Matriz de Coordenadas Principais*.

Ao fazermos o ajuste dos dados sobre estes pontos transformados, deveremos refletir a distribuição apresentada nos dados, por isto, o natural será considerar cada ponto com um peso igual à massa p_i , ($i = 1, \dots, I$). Logo, para todo α ($\alpha = 1, \dots, J$), a inércia dos pontos projeção da nuvem sobre $\mathbf{v}_\alpha$ pode ser expressada por:

$$\begin{aligned} \mathbf{F}_{(\alpha)}' \mathbf{D}_r \mathbf{F}_{(\alpha)} = \max \{ & \mathbf{f}(\mathbf{v})' \mathbf{D}_r \mathbf{f}(\mathbf{v}) \\ & : \mathbf{v} \in \mathbb{R}^J, \mathbf{v}' \mathbf{v} = 1, \mathbf{v}' \mathbf{v}_{\alpha-1} = 0, \dots, \mathbf{v}' \mathbf{v}_1 = 0 \}, \end{aligned} \quad (25)$$

equivalentemente, para as expressões respectivas à direita e à esquerda da igualdade acima, substituindo sucessivamente as expressões de $\mathbf{F}$ em (24), de $\mathbf{f}(\mathbf{v})$ em (23) e de $\mathbf{R}$ em (7), obtemos:

$$[\mathbf{X} \mathbf{v}_a]' [\mathbf{X} \mathbf{v}_a] = \max \{ [\mathbf{X} \mathbf{v}]' [\mathbf{X} \mathbf{v}] \quad (26)$$

$$: \mathbf{v} \in \mathbb{R}^J, \mathbf{v}'\mathbf{v}=1, \mathbf{v}'\mathbf{v}_{a-1}=0, \dots, \mathbf{v}'\mathbf{v}_1=0 \},$$

onde:

$$\mathbf{X} = \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} - \mathbf{r}\mathbf{c}') \mathbf{D}_c^{-\frac{1}{2}}. \quad (27)$$

Isto, equivale a buscar os autovalores e autovetores da matriz:

$$\mathbf{Z} = \mathbf{X}'\mathbf{X}, \quad (28)$$

(Veja demonstração no apêndice B). Logo, a matriz $\mathbf{V}$, com $\mathbf{V}'\mathbf{V} = \mathbf{I}$ é a matriz de autovetores de $\mathbf{Z}$ associados aos autovalores $\mu_1^2 \geq \mu_2^2 \geq \dots \geq \mu_J^2 \geq 0$ respectivamente i.e.:

$$\mathbf{Z} \mathbf{V} = \mathbf{V} \mathbf{D}_\mu^2, \quad (29)$$

onde $\mathbf{D}_\mu = \text{diag}\{\boldsymbol{\mu}\}$, com $\boldsymbol{\mu} = (\mu_1, \dots, \mu_J)$, e onde sem perda de generalidade, restringiremos $\boldsymbol{\mu}$ tal que $\mu_\alpha \geq 0$ ($\alpha = 1, \dots, J$) (Notemos que: para todo $\mathbf{v} \in \mathbb{R}^J$, o valor $[\mathbf{X} \mathbf{v}]' [\mathbf{X} \mathbf{v}]$ é sempre não negativo, pois trata-se de uma soma de quadrados. Isto justifica a escolha da apresentação dos autovalores como quadrados).

Observemos que os valores máximos procurados nas expressões (25) e (26), são exatamente os correspondentes autovalores μ_α^2 ($\alpha = 1, \dots, J$), pois de (29), temos que: $\mathbf{v}_\alpha' \mathbf{X}' \mathbf{X} \mathbf{v}_\alpha = \mu_\alpha^2$.

Assim temos que, os autovetores de $\mathbf{Z}$ definem os eixos principais dos perfis linha, e que o sub-espaco d -dimensional ($1 \leq d \leq J$) que melhor se ajusta aos dados em termos de distâncias quadradas ponderadas pelas massas, está definido pelos d autovetores $\mathbf{v}_d$, correspondentes aos d maiores autovalores de $\mathbf{Z}$ (Eckart e Young, 1936).

Simplificações

A seguir, veremos que existem algumas relações que permitem reduzir os cálculos. Inclusive, estas podem ser deduzidas diretamente da geometria exposta na seção 2.2.7-Perfis Transformados:

- a. O vetor $\mathbf{v}_J$, diretor do J -ésimo eixo, é igual a $(\mathbf{D}_c^{-\frac{1}{2}} \mathbf{1}) = (p_{.1}^{-\frac{1}{2}}, p_{.2}^{-\frac{1}{2}}, \dots, p_{.J}^{-\frac{1}{2}})'$, e é autovetor da matriz $\mathbf{Z}$, associado ao autovalor 0.

Com efeito:

$$\begin{aligned} \mathbf{X} \mathbf{v}_J &= [\mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{D}_c^{-\frac{1}{2}}] [\mathbf{D}_c^{\frac{1}{2}} \mathbf{1}] \\ &= \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{1} \\ &= \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} \mathbf{1} - \mathbf{r} \mathbf{c}' \mathbf{1}) \\ &= \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{r} - \mathbf{r}) \\ &= \mathbf{0}, \quad \text{onde } \mathbf{0} = (0, \dots, 0)' \in \mathbb{R}^I, \end{aligned}$$

com este resultado, temos que $0 = [\mathbf{X} \mathbf{v}_J]' [\mathbf{X} \mathbf{v}_J] = \mathbf{v}_J' \mathbf{Z} \mathbf{v}_J$, e portanto, $\mathbf{Z} \mathbf{v}_J = 0 \mathbf{v}_J$. ■

- b. O vetor $\mathbf{v}_J$ é também autovetor associado ao autovalor 1 de:

$$\mathbf{Z}^{\dagger} = \mathbf{X}^{\dagger} \mathbf{X}^{\dagger}, \quad (30)$$

$$\text{onde:} \quad \mathbf{X}^{\dagger} = \mathbf{D}_r^{-\frac{1}{2}} \mathbf{P} \mathbf{D}_c^{-\frac{1}{2}}, \quad (31)$$

Com efeito:

$$\begin{aligned} \mathbf{X}^{\dagger} \mathbf{v}_J &= [\mathbf{D}_r^{-\frac{1}{2}} \mathbf{P} \mathbf{D}_c^{-\frac{1}{2}}] [\mathbf{D}_c^{\frac{1}{2}} \mathbf{1}] \\ &= \mathbf{D}_r^{-\frac{1}{2}} \mathbf{P} \mathbf{1} \\ &= \mathbf{D}_r^{-\frac{1}{2}} \mathbf{r} \\ &= \mathbf{D}_r^{\frac{1}{2}} \mathbf{1} \\ &= \mathbf{v}_J \end{aligned}$$

com este resultado, temos que $\mathbf{v}_j' \mathbf{Z}' \mathbf{v}_j = [\mathbf{X}' \mathbf{v}_j]' [\mathbf{X}' \mathbf{v}_j] = \mathbf{v}_j' \mathbf{v}_j = 1$, e portanto, $\mathbf{Z}' \mathbf{v}_j = 1 \mathbf{v}_j$. ■

c. Todos os autovetores de $\mathbf{Z}$ distintos de $\mathbf{v}_j$, são também autovetores de $\mathbf{Z}'$.

Com efeito, os restantes autovetores de $\mathbf{Z}$ são ortogonais a $\mathbf{v}_j$, ou seja, para $\alpha = 1, \dots, J-1$, $(\mathbf{D}_c^{\frac{1}{2}} \mathbf{1})' \mathbf{v}_\alpha = 0$. Com esta relação, deduzimos que, para todo $\alpha = 1, \dots, J-1$, temos:

$$\begin{aligned} \mathbf{X} \mathbf{v}_\alpha &= [\mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{D}_c^{-\frac{1}{2}}] \mathbf{v}_\alpha \\ &= \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} - \mathbf{r} \mathbf{c}') \mathbf{D}_c^{-\frac{1}{2}} \mathbf{v}_\alpha \\ &= \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} \mathbf{D}_c^{-\frac{1}{2}} \mathbf{v}_\alpha - \mathbf{r} \mathbf{c}' \mathbf{D}_c^{-\frac{1}{2}} \mathbf{v}_\alpha) \\ &= \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} \mathbf{D}_c^{-\frac{1}{2}} \mathbf{v}_\alpha - \mathbf{r} (\mathbf{D}_c^{\frac{1}{2}} \mathbf{1})' \mathbf{v}_\alpha) \\ &= \mathbf{D}_r^{-\frac{1}{2}} \mathbf{P} \mathbf{D}_c^{-\frac{1}{2}} \mathbf{v}_\alpha \\ &= \mathbf{X}' \mathbf{v}_\alpha, \end{aligned} \tag{32}$$

com este resultado, temos que $[\mathbf{X}' \mathbf{v}_\alpha]' [\mathbf{X}' \mathbf{v}_\alpha] = [\mathbf{X} \mathbf{v}_\alpha]' [\mathbf{X} \mathbf{v}_\alpha] = \mu_\alpha^2$, e portanto, $\mathbf{Z}' \mathbf{v}_\alpha = \mathbf{Z} \mathbf{v}_\alpha = \mu_\alpha^2 \mathbf{v}_\alpha$. ■

d. Para $\alpha = 1, \dots, J-1$, temos que $\mu_\alpha \leq 1$.

A prova desta afirmação será feita mais adiante.

Vemos então que, a decomposição da matriz $\mathbf{Z}'$, contém a decomposição da matriz $\mathbf{Z}$, i.e.:

$$\mathbf{Z}' \mathbf{V}^* = \mathbf{V}^* \mathbf{D}_{\mu^*}^2, \tag{33}$$

onde:

$$\begin{aligned} \mathbf{V}^* &= [(\mathbf{D}_c^{\frac{1}{2}} \mathbf{1}) \mid \mathbf{v}_1 \mid \dots \mid \mathbf{v}_{J-1}] \text{ e} \\ \mathbf{D}_{\mu^*}^2 &= \text{diag}\{\boldsymbol{\mu}^*\}, \quad \text{com } \boldsymbol{\mu}^* = (1, \mu_1, \dots, \mu_{J-1})'. \end{aligned}$$

Conseqüências

- Na AC, a centralização dos perfis é uma operação que remove exatamente uma dimensão do espaço original.

- Os pontos se encontram num sub-espaco de dimensao D , onde $0 \leq D \leq J-1$. Ainda, se estivessemos no caso em que $I < J$, teriamos que $1 \leq D \leq I-1$, pois só temos I vetores originais provenientes das linhas de $\mathbf{R}$, que no máximo formam um sub-espaco I -dimensional, os quais centralizados, perdem uma dimensao, como acabamos de ver.
- A análise pode ser realizada sobre a matriz $\mathbf{Z}^*$ (não centrada), com a condicao de desprezar o autovalor 1 e seu autovetor correspondente.
- Utilizando a relacao em (32), também poderemos simplificar os cálculos das coordenadas principais dos pontos linha sobre o α -ésimo eixo ($\alpha = 1, \dots, J-1$). Assim, partindo de (24) e (27), obtemos uma nova expressao para a matriz de coordenadas principais das linhas:

$$\mathbf{F} = \mathbf{D}_r^{-\frac{1}{2}} \mathbf{X}^* \mathbf{V}, \quad \text{ou} \quad \mathbf{F} = \mathbf{R} \mathbf{D}_c^{-\frac{1}{2}} \mathbf{V}. \quad (34)$$

2.3.2 A Análise em $\mathbb{R}^I$

Procedendo com a análise de forma análoga e simétrica como em $\mathbb{R}^J$, nos limitaremos a dar os resultados.

O Problema Proposto

O objetivo é obter uma representacao simplificada dos perfis coluna, dando uma ideia das distancias $\mathbf{X}^2$ entre eles, portanto, representaremos os perfis da matriz $\mathbf{C} \mathbf{D}_r^{-\frac{1}{2}}$.

A Solução

Como queremos encontrar uma base $\mathbf{U} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_I\}$ de $\mathbb{R}^I$, tal que:

Para todo β ($\beta = 1, \dots, I$), a inercia dos pontos projeção da nuvem sobre $\mathbf{u}_\beta$ seja máxima, $\mathbf{u}_\beta' \mathbf{u}_\beta = 1$ e $\mathbf{u}_\beta' \mathbf{u}_{\beta-1} = 0, \dots, \mathbf{u}_\beta' \mathbf{u}_1 = 0$.

Definiremos a *Matriz de Coordenadas Principais* para as colunas sobre os eixos procurados $\mathbf{u}_1, \dots, \mathbf{u}_\beta, \dots, \mathbf{u}_I$:

$$\mathbf{G} = (\mathbf{C} - \mathbf{1}\mathbf{r}') \mathbf{D}_r^{-1} \mathbf{U}, \quad (35)$$

onde $\mathbf{U} = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \dots \mid \mathbf{u}_I]$, com $\mathbf{U}'\mathbf{U} = \mathbf{I}$ e

$$\mathbf{G} = [\mathbf{g}_{(1)} \mid \mathbf{g}_{(2)} \mid \dots \mid \mathbf{g}_{(I)}].$$

O problema se reduz a buscar os autovalores e autovetores da matriz:

$$\mathbf{W} = \mathbf{X} \mathbf{X}', \quad (36)$$

Encontramos que os vetores $\mathbf{u}_1, \dots, \mathbf{u}_I$ são os autovetores da matriz de $\mathbf{W}$ associados aos autovalores $\tau_1^2 \geq \tau_2^2 \geq \dots \geq \tau_I^2 = 0$ respectivamente. i.e.:

$$\mathbf{W} \mathbf{U} = \mathbf{U} \mathbf{D}_\tau^2, \quad (37)$$

onde $\mathbf{D}_\tau = \text{diag}\{\tau\}$, com $\tau = (\tau_1, \dots, \tau_I)$ e $1 \geq \tau_1 \geq \tau_2 \geq \dots \geq \tau_I = 0$.

Os autovetores de $\mathbf{W}$ definem os eixos principais dos perfis coluna, e a inércia contabilizada por eles são os correspondentes autovalores τ_β^2 ($\beta = 1, \dots, I$).

O sub-espço d -dimensional ($1 \leq d \leq I$) que melhor se ajusta aos dados em termos de distâncias quadradas ponderadas pelas massas, está definido pelos d autovetores $\mathbf{u}_\alpha$, correspondentes aos d maiores autovalores de $\mathbf{W}$.

Simplificações

A decomposição de $\mathbf{W}^* = \mathbf{X}^* \mathbf{X}'^*$, contém a decomposição de $\mathbf{W}$, i.e.:

$$\mathbf{W}^* \mathbf{U}^* = \mathbf{U}^* \mathbf{D}_{\tau^*}^2, \quad (38)$$

onde:

$$\mathbf{U}^* = [(\mathbf{D}_r^{-\frac{1}{2}} \mathbf{1}) \mid \mathbf{u}_1 \mid \dots \mid \mathbf{u}_{I-1}] \text{ e}$$

$$\mathbf{D}_{\tau^*} = \text{diag}\{\tau^*\}, \quad \text{com } \tau^* = (1, \tau_1, \dots, \tau_{I-1})'.$$

Conseqüências

- A centralização dos perfis, é uma operação que remove exatamente uma dimensão do espaço original.
- Os pontos se encontram num sub-espaco de dimensão D , onde $0 \leq D \leq I-1$. Adicionalmente, se estivermos no caso em que $J < I$, teríamos que $1 \leq D \leq J-1$, pois só temos J vetores originais provenientes das linhas de $\mathbf{C}$, que no máximo formam um sub-espaco J -dimensional, os quais centralizados, perdem uma dimensão.
- A análise pode ser realizada sobre a matriz $\mathbf{W}^*$, com a condição de desprezar o autovalor 1 e seu autovetor correspondente.
- Podemos simplificar os cálculos para a matriz de coordenadas principais das colunas:

$$\mathbf{G} = \mathbf{D}_c^{-\frac{1}{2}} \mathbf{X}^* \mathbf{U}, \quad \text{ou} \quad \mathbf{G} = \mathbf{C} \mathbf{D}_r^{-\frac{1}{2}} \mathbf{U}. \quad (39)$$

2.3.3 Relações entre os Espaços

A seguir, veremos as relações existentes entre os autovalores de $\mathbf{Z} = \mathbf{X}'\mathbf{X}$ e $\mathbf{W} = \mathbf{X}\mathbf{X}'$, com as quais só precisaremos fazer uma decomposição em valores singulares: a da matriz $\mathbf{X}$ (as relações estão expostas no apêndice seção C).

- a. Existe um D , onde $0 \leq D \leq \min\{I, J\} - 1$, tal que os autovalores $\tau_{D+1} = \dots = \tau_I = \mu_{D+1} = \dots = \mu_J = 0$.
- b. Se $D > 0$, temos que para $\alpha = 1, \dots, D$, os correspondentes autovalores μ_α^2 de $\mathbf{Z}$ e τ_α^2 de $\mathbf{W}$, são iguais, i.e.: $\mu_\alpha = \tau_\alpha$.
Doravante, para evitar duplicidade na notação, só usaremos μ .

c. A matriz $\mathbf{X}$ decomposta em valores singulares:

$$\mathbf{X} = \mathbf{U} \mathbf{D}_\mu \mathbf{V}', \quad (40)$$

onde, tomando em conta os resultados em (a.) e (b.), daqui em diante só consideraremos $\mathbf{U}$, $\mathbf{V}$ e $\mathbf{D}_\mu$, com suas dimensões não nulas, i.e.:

$$\begin{aligned} \mathbf{U} &= [\mathbf{u}_1 \mid \dots \mid \mathbf{u}_D], \text{ com } \mathbf{U}'\mathbf{U} = \mathbf{I}_I, \\ \mathbf{V} &= [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_D], \text{ com } \mathbf{V}'\mathbf{V} = \mathbf{I}_J \text{ e} \\ \mathbf{D}_\mu &= \text{diag}\{\boldsymbol{\mu}\}, \quad \text{com } \boldsymbol{\mu} = (\mu_1, \dots, \mu_D)' \\ &\quad \text{e } \mu_1 \geq \dots \geq \mu_D > 0. \end{aligned}$$

onde $\mathbf{u}_\alpha$ e $\mathbf{v}_\alpha$ ($\alpha = 1, \dots, D$), são os mesmos que achamos nas seções anteriores.

Recordando as simplificações feitas nas seções anteriores, adicionaremos que, no caso particular da AC, a centralização dos perfis, é uma operação simétrica que remove exatamente uma dimensão de cada espaço original; pois a decomposição em valores singulares da matriz $\mathbf{X}^* = \mathbf{D}_r^{-\frac{1}{2}} \mathbf{P} \mathbf{D}_c^{-\frac{1}{2}}$, contém a decomposição em valores singulares da matriz $\mathbf{X} = \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P} - \mathbf{r}\mathbf{c}') \mathbf{D}_c^{-\frac{1}{2}}$, i.e.:

$$\mathbf{X}^* = \mathbf{U}^* \mathbf{D}_{\mu^*} \mathbf{V}^{*'}, \quad (41)$$

onde, de agora em diante só consideraremos $\mathbf{U}^*$, $\mathbf{V}^*$ e $\mathbf{D}_{\mu^*}$, com suas dimensões não nulas, i.e.:

$$\begin{aligned} \mathbf{U}^* &= [(\mathbf{D}_r^{-\frac{1}{2}} \mathbf{1}) \mid \mathbf{u}_1 \mid \dots \mid \mathbf{u}_D], \\ \mathbf{V}^* &= [(\mathbf{D}_c^{-\frac{1}{2}} \mathbf{1}) \mid \mathbf{v}_1 \mid \dots \mid \mathbf{v}_D] \text{ e} \\ \mathbf{D}_{\mu^*} &= \text{diag}\{\boldsymbol{\mu}^*\}, \quad \text{com } \boldsymbol{\mu}^* = (1, \mu_1, \dots, \mu_D). \\ &\quad \text{e } 1 \geq \mu_1 \geq \dots \geq \mu_D > 0. \end{aligned}$$

Naturalmente, visto que usaremos somente as dimensões não nulas, passaremos a considerar as matrizes de coordenadas principais da seguinte maneira:

$$\mathbf{F} = [\mathbf{F}_1 \mid \dots \mid \mathbf{F}_D] \quad \text{e} \quad \mathbf{G} = [\mathbf{G}_1 \mid \dots \mid \mathbf{G}_D], \quad (42)$$

já que para $\alpha > D$, temos que: $\mathbf{F}_\alpha = \mathbf{0}_I$ e $\mathbf{G}_\alpha = \mathbf{0}_J$.

d. As matrizes $\mathbf{V}$ e $\mathbf{U}$, estão ligadas pelas seguintes relações:

$$\mathbf{U} = \mathbf{X} \mathbf{V} \mathbf{D}_\mu^{-1} \quad \text{e} \quad \mathbf{V} = \mathbf{X}' \mathbf{U} \mathbf{D}_\mu^{-1}, \quad (43)$$

daqui, observemos a existência de uma proporcionalidade entre os eixos de um espaço e do outro.

Equações de Transição

As relações entre os dois espaços, indicam que há proporcionalidade entre os vetores correspondentes, e, por conseguinte, também haverá proporcionalidade entre as coordenadas dos pontos correspondentes aos eixos. De modo que, nos permitirão representar simultaneamente os pontos linha e coluna sobre os mesmos gráficos, enriquecendo a interpretação.

As coordenadas principais (não nulas) cumprem as seguintes equações:

$$\mathbf{F} = \mathbf{R} \mathbf{G} \mathbf{D}_\mu^{-1} \quad \text{e} \quad \mathbf{G} = \mathbf{C} \mathbf{F} \mathbf{D}_\mu^{-1}. \quad (44)$$

Alguns autores chamam estas de *Equações de Transição* (Greenacre, 1984; Lebart et al., 1984), pois nos permitem transitar de um espaço para outro. Outros autores preferem o termo *Relações Baricêntricas* (Grande E. e Abascal F., 1989).

Prova:

De (43), temos:

$$\mathbf{U} \mathbf{D}_\mu = \mathbf{X} \mathbf{V} \quad \text{e} \quad \mathbf{V} \mathbf{D}_\mu = \mathbf{X}' \mathbf{U},$$

logo, pré-multiplicando estas equações por $\mathbf{D}_r^{-1}$ e $\mathbf{D}_c^{-1}$ respectivamente, obtemos que são iguais a $\mathbf{F}$ e $\mathbf{G}$ respectivamente, i.e.:

$$\mathbf{F} = \mathbf{D}_r^{-1} \mathbf{U} \mathbf{D}_\mu \quad \text{e} \quad \mathbf{G} = \mathbf{D}_c^{-1} \mathbf{V} \mathbf{D}_\mu, \quad (45)$$

ou equivalentemente:

$$\mathbf{U} = \mathbf{D}_r \mathbf{F} \mathbf{D}_\mu^{-1} \quad \text{e} \quad \mathbf{V} = \mathbf{D}_c \mathbf{G} \mathbf{D}_\mu^{-1}, \quad (46)$$

logo, substituindo estas expressões de $\mathbf{V}$ e $\mathbf{U}$ respectivamente em (34) e (39), temos que os conjuntos das coordenadas efetivas $\mathbf{F}$ e $\mathbf{G}$ podem ser expressadas como:

$$\mathbf{F} = \mathbf{R} \mathbf{G} \mathbf{D}_\mu^{-1} \quad \text{e} \quad \mathbf{G} = \mathbf{C} \mathbf{F} \mathbf{D}_\mu^{-1},$$

que é o resultado que queríamos.■

Conclusões

Visto que ambos os conjuntos de pontos estão dentro de subespaços da mesma dimensão, e se colocarmos superpostas as dimensões correspondentes, ao reescrevermos as identidades em (44):

$$f_{ia} = 1/\mu_a \sum_j (p_{ij}/p_{i.}) g_{ja} \quad \text{e} \quad g_{ja} = 1/\mu_a \sum_i (p_{ij}/p_{.j}) f_{ia}$$

onde $\sum_j (p_{ij}/p_{i.}) = 1$, e $\sum_i (p_{ij}/p_{.j}) = 1$, veremos que:

- A projeção de uma linha i ($i = 1, \dots, I$) sobre o α -ésimo eixo ($\alpha = 1, \dots, D$), f_{ia} , é o baricentro, salvo um coeficiente de dilatação ($1/\mu_a$), das projeções das J colunas sobre o mesmo eixo, onde cada coluna está afetada por um peso ($p_{ij}/p_{i.}$), que é a sua importância relativa em i .
- A projeção de uma coluna j ($j = 1, \dots, J$) sobre o α -ésimo eixo ($\alpha = 1, \dots, D$), g_{ja} , é o baricentro, salvo um coeficiente de dilatação ($1/\mu_a$), das projeções das I linhas sobre o mesmo eixo,

onde cada linha está afetada por um peso $(p_{ij}/p_{.j})$, que é a sua importância relativa em j .

Neste contexto, os perfis são usados como pesos para o cálculo dos baricentros. Estes resultados, nos permitem *representar simultaneamente* as linhas e colunas.

Consequência

Das fórmulas baricêntricas, se deduz que, os valores próprios μ_α^2 são menores ou iguais que a unidade. Em outras palavras, sem perda de generalidade:

$$0 < \mu_\alpha \leq 1 \quad , \text{ para todo } \alpha = 1, \dots, D$$

Tal afirmação, será demonstrada por redução ao absurdo:

Suponhamos que existe um α ($\alpha = 1, \dots, D$) tal que $\mu_\alpha > 1$.

Logo, $1/\mu_\alpha < 1$.

Assim, da primeira das relações, obteríamos para todo i ($i = 1, \dots, I$):

$$\begin{aligned} |f_{i\alpha}| &= 1/\mu_\alpha \left| \sum_j (p_{ij}/p_{.j}) g_{j\alpha} \right| < \left| \sum_j (p_{ij}/p_{.j}) g_{j\alpha} \right| \\ &\leq \sum_j (p_{ij}/p_{.j}) |g_{j\alpha}| \leq \max_j |g_{j\alpha}|, \end{aligned}$$

em particular, isto se cumpre para:

$$\max_i |f_{i\alpha}| < \max_j |g_{j\alpha}|. \quad (a)$$

Por outro lado, a partir da segunda das relações, obteríamos para todo j ($j = 1, \dots, J$):

$$\begin{aligned} |g_{j\alpha}| &= 1/\mu_\alpha \left| \sum_i (p_{ij}/p_{.j}) f_{i\alpha} \right| < \left| \sum_i (p_{ij}/p_{.j}) f_{i\alpha} \right| \\ &\leq \sum_i (p_{ij}/p_{.j}) |f_{i\alpha}| \leq \max_i |f_{i\alpha}|, \end{aligned}$$

em particular:

$$\max_j |g_{j\alpha}| < \max_j |f_{ja}|. \quad (b)$$

Não podemos obter os resultados (a) e (b) simultaneamente, pois são contraditórios.

Visto que a nossa suposição nos levará a obter uma contradição, só podemos colocar como verdadeira a negação desta, i.e.: só pode acontecer que:

$$0 < \mu_\alpha \leq 1, \text{ para todo } \alpha (\alpha = 1, \dots, D) . \blacksquare$$

2.3.4 Reconstrução da Tabela de Contingência

É possível reconstruir a tabela $\mathbf{P}$ a partir das coordenadas principais encontradas. Com efeito, partindo de (40), e substituindo $\mathbf{X}$, $\mathbf{U}$ e $\mathbf{V}$ com suas expressões dadas em (27) e (46) respectivamente, temos que:

$$\mathbf{P} = \mathbf{rc}' + \mathbf{D}_r \mathbf{F} \mathbf{D}_\mu^{-1} \mathbf{G}' \mathbf{D}_c. \quad (47)$$

A partir desta equação, claramente podemos ver que AC decompõe os resíduos da esperança de $\mathbf{P}$ sob hipótese de independência.

Na prática, acontece freqüentemente que os dois ou três primeiros eixos recolhem a maior quantidade de informação. Logo, se são representados os pontos sobre o plano formado pelos dois primeiros eixos, é possível obter uma boa visão dos dados. Esta representação será tanto mais fiel, quanto maior for a percentagem de inércia explicada pelos primeiros eixos. Para obter maior precisão na imagem, se completa com os eixos terceiro, quarto, etc...

Para $d = 1, \dots, D$, definiremos a *Matriz de Reconstrução da Tabela de ordem d*:

$$\mathbf{P}^{(d)} = \mathbf{rc}' + \mathbf{D}_r \mathbf{F}^{(d)} \mathbf{D}_{\mu^{(d)}}^{-1} \mathbf{G}^{(d)'} \mathbf{D}_c \quad (48)$$

onde:

$$\mathbf{F}^{(d)} = [\mathbf{F}_1 | \dots | \mathbf{F}_d] \quad , \quad \mathbf{G}^{(d)} = [\mathbf{G}_1 | \dots | \mathbf{G}_d] \quad , \text{ e}$$

$$\mathbf{D}_{\mu}^{(d)} = \text{diag}\{\boldsymbol{\mu}^{(d)}\} \quad , \quad \text{com } \boldsymbol{\mu}^{(d)} = (\mu_1, \dots, \mu_d).$$

Quando os $D - d$ menores valores singulares são suficientemente pequenos, $\mathbf{P}^{(d)}$ pode aproximar-se bastante de $\mathbf{P}$, um resultado interessante é que a aproximação em $\mathbf{P}^{(d)}$ aos dados originais em $\mathbf{P}$, pode ser avaliada da seguinte maneira:

observemos que, para $\mathbf{X}^{(d)} = \mathbf{D}_r^{-\frac{1}{2}} (\mathbf{P}^{(d)} - \mathbf{r}\mathbf{c}') \mathbf{D}_c^{-\frac{1}{2}}$:

$$\mathbf{X}^{(d)} = \mathbf{D}_r^{\frac{1}{2}} \mathbf{F}^{(d)} \mathbf{D}_{\mu}^{(d)-1} \mathbf{G}^{(d)} \mathbf{D}_c^{\frac{1}{2}}, \quad (49)$$

podemos calcular:

$$\tau^{(d)} = \frac{\sum_{ij} (X^{(d)}_{ij})^2}{\sum_{ij} (X_{ij})^2}, \quad (50)$$

isto é, a soma das normas quadradas dos pontos de $\mathbf{X}^{(d)}$, dividida pela soma das normas quadradas dos pontos de $\mathbf{X}$, que resulta em:

$$\tau^{(d)} = \frac{\text{tr} [\mathbf{X}^{(d)} , \mathbf{X}^{(d)}]}{\text{tr} [\mathbf{X}' \mathbf{X}]} = \frac{\sum_a^d \mu_a^2}{\sum_a^D \mu_a^2} \quad (51)$$

comparando este resultado com (17), isto significa que $\tau^{(d)}$ é a porção da inércia total contabilizada pelos d primeiros eixos, é uma medida da qualidade da recuperação dos dados em $\mathbf{P}^{(d)}$.

2.4 Interpretação dos Resultados

Como vimos anteriormente, podemos resumir a AC como um modelo simples, o modelo de independência é eleito, e os resíduos deste modelo são decompostos e graficados. É desta forma que obtemos a figura da interação entre linhas e colunas.

Nesta seção, veremos a distribuição das coordenadas principais, a interpretação da posição dos pontos tanto com relação aos pontos da mesma nuvem, como aos pontos da outra nuvem. Logo, a definição das coordenadas estandardizadas, e suas características. Também veremos os auxílios para interpretar os resultados, construídas a partir da decomposição da inércia. Estes auxílios são necessários para interpretar os eixos, e sobretudo, quando não podemos graficar os pontos completamente (i.e. os pontos estão num espaço de dimensão maior que 2 ou 3). Finalmente, veremos o uso de elementos suplementares, que poderiam ser úteis na interpretação.

2.4.1 Interpretação dos Pontos

Em Coordenadas Principais

Distribuição das Coordenadas Principais

Com respeito aos eixos principais, as nuvens têm centróides na origem, i.e.:

$$\mathbf{F}'\mathbf{r} = \mathbf{0}_J \qquad \mathbf{G}'\mathbf{c} = \mathbf{0}_I, \qquad (52)$$

estes resultados são facilmente verificáveis, ao substituir as definições para $\mathbf{F}$ e $\mathbf{G}$ em (24) e (35) respectivamente.

A Inércia (ou variância total ponderada pela massa) de cada nuvem ao longo do α -ésimo eixo ($\alpha = 1, \dots, D-1$), é igual a μ_α^2 , e a chamaremos de α -ésima *Inércia Principal*. Os produtos cruzados, ponderados pela massa se anulam como esperado pela construção da ortogonalidade entre os eixos em (24) e (35), i.e.:

$$\mathbf{F}' \mathbf{D}_r \mathbf{F} = \mathbf{D}_\mu^2 \quad \text{e} \quad \mathbf{G}' \mathbf{D}_c \mathbf{G} = \mathbf{D}_\mu^2. \quad (53)$$

para obter estes resultados, bastará substituir na direita destas equações, expressões para $\mathbf{F}$ e $\mathbf{G}$ em (34) e (39) respectivamente, e logo usar a decomposição em (41).

Observemos que para todo α ($\alpha = 1, \dots, D$) μ_α é uma medida de correlação entre as coordenadas principais f_{ia} e g_{ja} ($i = 1, \dots, I$, $j = 1, \dots, J$), pois de (47) com (52) e (53), se cumpre:

$$(\mathbf{F} \mathbf{D}_\mu^{-1})' \mathbf{P} (\mathbf{G} \mathbf{D}_\mu^{-1}) = \mathbf{D}_\mu. \quad (54)$$

Representação em Coordenadas Principais

Quando representarmos os pontos linha em coordenadas principais, a distância entre eles é a distância $\mathbf{X}^2$, logo com relação as posições entre eles, temos os seguintes resultados:

- a. Se duas linhas têm uma estrutura semelhante (i.e. seus perfis são similares), ambas as linhas relacionam-se de modo similar com as colunas. Logo, sua situação será próxima no gráfico. Mas cabe advertir que, a inversa não é sempre verdadeira; isto dependerá da qualidade de representação dos pontos, que explicaremos mais a frente.
- b. Quando duas linhas estão muito distanciadas, estão relacionadas de modos diferentes com as colunas. Quando dois pontos linha estão em direções opostas do centro, eles desviam de formas opostas do perfil médio (por exemplo, duas linhas são opostas se, dada a primeira linha, uma coluna tem proporção baixa, dada a segunda linha a coluna anterior tem proporção alta).

Quando representarmos os pontos coluna em coordenadas principais, os resultados serão análogos aos anteriores.

Quando representarmos as linhas e colunas simultaneamente sobre o mesmo gráfico em termos das coordenadas principais, devemos considerar as equações de transição em (44). Reescrevendo a primeira delas, temos:

$$\mathbf{F}_i = \mathbf{D}_\mu^{-1} \{ \sum_j (p_{ij}/p_{i.}) \mathbf{G}_j \}$$

onde $\mathbf{F}_i = (f_{i1}, \dots, f_{iD})'$, $\mathbf{G}_j = (g_{j1}, \dots, g_{jD})'$ e $\sum_j (p_{ij}/p_{i.}) = 1$.

Graças a esta, podemos interpretar as posições linhas com relação as posições das colunas, o seguinte:

- a. Dada a linha i , quanto maior seja a proporção de uma coluna j , esta linha será mais atraída na direção da coluna j . Pois, o vetor resultante $\mathbf{F}_i$ contém uma proporção maior do vetor $\mathbf{D}_\mu^{-1} \mathbf{G}_j$ (que terá quase a mesma direção e é uma dilatação do vetor $\mathbf{G}_j$). Em particular, se um perfil linha tem a forma de $\mathbf{e}_j$ ($j = 1, \dots, J$), onde $\mathbf{e}_j$ é o j -ésimo vetor da base canônica de $\mathbb{R}^J$, a representação deste se sobreporá com o vetor $\mathbf{D}_\mu^{-1} \mathbf{G}_j$ (e em geral, se localizará na mesma direção de $\mathbf{G}_j$).
- b. A proximidade de uma linha i a uma coluna j só pode ser interpretada como associadas, se ambos os pontos estiverem longe da origem, na periferia da nuvem. Pois, caso contrário, consideremos duas situações que podem acontecer: 1º) Quando uma linha tem um perfil próximo ao perfil médio, se encontrará 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 (i.e. dado i , para todo j , $p_{ij}/p_{i.} - p_{.j}$ é pequeno, isto é equivalente a p_{ij} estar próximo a $p_{i.} p_{.j}$, e indica que a casela (i, j) não contribui a desviação da independência). 2º) A proximidade entre i e j poderia dever-se não a que j tem importância em i , e sim, que os que têm importância sobre i são outras colunas j' e j'' que se encontram muito longe de j e em extremos opostos.

- c. Por tanto, a posição de um ponto linha deve ser interpretada com relação as posições de todos os pontos coluna.

Agora, para interpretar as posições das colunas com relação às linhas, nos basearemos na segunda equação de transição em (44) e os resultados serão análogos aos anteriores.

Em Coordenadas Estandarizadas

A representação dos pontos em coordenadas principais, é uma representação natural, imposta por nossa definição das duas geometrias simétricas e duais. Contudo, haverá algumas situações nas quais estaremos dispostos a sacrificar a natureza destas com o propósito de ganhar outras vantagens, estandardizando as coordenadas principais.

Faremos uma transformação das coordenadas principais, de modo que possamos obter, ao longo de cada eixo, inércias iguais a um. A estas novas coordenadas as chamaremos de *Coordenadas Estandarizadas*, para que sejam distinguidas das coordenadas principais e as denotaremos por Φ e Γ para cada espaço respectivamente. Assim:

$$\Phi = F D_{\mu}^{-1} \quad \text{e} \quad \Gamma = G D_{\mu}^{-1}, \quad (55)$$

Estas matrizes são chamadas por alguns autores de Fatores (Lebart, Morineau e Warwick, 1984).

É fácil comprovar de (52), que com respeito aos eixos estandardizados, as nuvens têm centróides na origem, i.e.:

$$\Phi' r = 0_J \quad \text{e} \quad \Gamma' c = 0_I, \quad (56)$$

de (53) temos que, as colunas de Φ e Γ estão estandardizadas segundo:

$$\Phi' D_r \Phi = I \quad \text{e} \quad \Gamma' D_c \Gamma = I, \quad (57)$$

isto significa que, para as coordenadas estandardizadas, obtemos variâncias unitárias ao longo dos eixos e estes são ortogonais entre si (claro, pois com a

transformação de (55), não houve rotação, somente uma dilatação das coordenadas desde que $1/\mu_\alpha \geq 1$).

De (54), temos que, para todo α ($\alpha = 1, \dots, D$) μ_α é uma medida de correlação entre as coordenadas estandardizadas $\Phi_{i\alpha}$ e $\Gamma_{j\alpha}$ desde que se cumpra:

$$\Phi' P \Gamma = D_\mu. \quad (58)$$

As equações de transição, conseguem a mesma forma para as coordenadas estandardizadas. Com efeito, bastará pós-multiplicar (44) por D_μ^{-1} , assim:

$$\Phi = R \Gamma D_\mu^{-1} \quad \text{e} \quad \Gamma = C \Phi D_\mu^{-1}. \quad (59)$$

De (47), podemos reconstruir a tabela a partir das coordenadas estandardizadas (ou fatores), segundo:

$$P = rc' + D_r \Phi D_\mu \Gamma' D_c. \quad (60)$$

Esta reconstrução, é conhecida como *Identidade de Fisher* (Lancaster, 1969), e significa simplesmente uma reparametrização da reconstrução em (47).

Goodman (1985a, 1986), identificou (60) com as restrições em (56) e (57), como um *Modelo de Correlação RC Saturado*. Logo, a reconstrução em (47), com as restrições em (52) e (53), pode ser vista como uma reparametrização do modelo de correlação RC saturado. A diferença é que neste modelo, as estimações dos parâmetros, geralmente são feitas por máxima verossimilhança e não como estabelece a AC, por quadrados mínimos.

Podemos reescrever a matriz $P^{(d)}$ de (48) a partir das coordenadas estandardizadas:

$$P^{(d)} = rc' + D_r \Phi^{(d)} D_{\mu^{(d)}} \Gamma^{(d)'} D_c \quad (61)$$

onde:

$$\Phi^{(d)} = [\phi_1 | \dots | \phi_d] \quad , \quad \Gamma^{(d)} = [\gamma_1 | \dots | \gamma_d] \quad .e$$

$$D_{\mu}^{(d)} = \text{diag}\{\mu^{(d)}\} \quad , \quad \text{com } \mu^{(d)} = (\mu_1, \dots, \mu_d).$$

A equação acima, corresponde ao *modelo de correlação RC não saturado*, definido por Goodman (1985a, 1986). Este modelo pode ser testado com $(I-d-1)(J-d-1)$ graus de liberdade (veja Goodman, 1986).

Representação em Coordenadas Estandarizadas

Se representarmos ambos os conjuntos com as coordenadas estandarizadas, nossa interpretação deverá ser feita de forma análoga como quando representamos os pontos com as coordenadas principais, pois os centróides, as equações de transição, e outros resultados, são análogos aos das coordenadas principais. A diferença é que quando as inércias principais são muito diferentes, a representação em coordenadas estandarizadas, em certa forma, dá "esfericidade" ao gráfico.

Relações entre os dois tipos de Coordenadas

Tomando em conta as transformações em (55), as fórmulas de transição entre as coordenadas dos conjuntos de pontos em (44) ficam, para **F** e **Γ** e para **G** e **Φ** :

$$\mathbf{F} = \mathbf{R} \mathbf{\Gamma} \quad e \quad \mathbf{G} = \mathbf{C} \mathbf{\Phi}, \quad (62)$$

Conclusões

- A projeção de uma linha i ($i = 1, \dots, I$) sobre o eixo α ($\alpha = 1, \dots, D$), $f_{i\alpha}$, é exatamente o baricentro das projeções Γ_j das J colunas sobre o mesmo eixo, onde cada coluna está afetada por um peso (p_{ij}/p_i) que é a sua importância relativa em i .
- A projeção de uma coluna j ($j = 1, \dots, J$) sobre o eixo α ($\alpha = 1, \dots, D$), $g_{j\alpha}$, é o baricentro das projeções ϕ_i das I linhas sobre o mesmo eixo, onde cada linha está afetada por um peso (p_{ij}/p_j) que é a sua importância relativa em j .

Estes resultados, nos permitem combinar os dois tipos de coordenadas para *representar simultaneamente* as linhas e colunas.

Representação Combinando Ambos os Tipos de Coordenadas

Quando representarmos nos gráficos, um dos conjuntos de pontos com as coordenadas principais, e o outro com as coordenadas estandardizadas, nossa interpretação pode ser grandemente facilitada. Nosso ganho ao combinar ambos os tipos de coordenadas, se baseia nas equações de transição expressadas em (62), que usaremos para interpretar as posições entre linhas e colunas.

Por exemplo, num gráfico onde as linhas estão representadas pelas coordenadas principais e as colunas estão representadas pelas coordenadas estandardizadas, 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 de (7). i.e. refazendo a expressão em (62), temos:

$$\mathbf{F}_i = \sum_j (p_{ij}/p_{i.}) \mathbf{r}_j,$$

onde $\mathbf{F}_i = (f_{i1}, \dots, f_{iD})'$ e $\mathbf{r}_j = (r_{j1}, \dots, r_{jD})'$ e $\sum_j (p_{ij}/p_{i.}) = 1$.

Assim, quando representarmos as categorias linha em coordenadas principais e as categorias colunas em coordenadas estandardizadas, veremos que, quanto maior for a proporção de uma coluna j dada a linha i , haverá mais proximidade de um ponto i a um ponto j . Pois, o vetor resultante $\mathbf{F}_i$ contém uma porção maior do vetor $\mathbf{r}_j$. Em particular, se um perfil linha tem a forma de $\mathbf{e}_j$ ($j = 1, \dots, J$), onde $\mathbf{e}_j$ é o j -ésimo vetor da base canônica de $\mathbb{R}^J$, a representação deste se sobreporá com o vetor $\mathbf{r}_j$. Logo, também podemos concluir que, devido à ponderação do espaço feita no início, isto também indica que quanto menor a massa da coluna j , mais longe se encontrará o ponto $\mathbf{r}_j$ do centróide.

Em geral, quanto maior for a diferença $(p_{ij}/p_{i.} - p_{.j})$, então os pontos i e j estarão mais perto. Recíprocamente, se para todo j ($j = 1, \dots, J$), a diferença $(p_{ij}/p_{i.} - p_{.j})$ for pequena, então i se localizará perto da origem.

Achamos que esta forma de representação pode ajudar a compreender melhor a representação quando ambos os conjuntos estão em coordenadas principais.

Ao representarmos as colunas em coordenadas principais e as linhas em coordenadas estandardizadas, todos estes resultados podem ser obtidos de forma análoga.

Para representar os pontos, existem outras alternativas, por exemplo, com coordenadas do tipo $\Phi^* = \mathbf{U}\mathbf{D}_\mu^{\frac{1}{2}}$ e $\Gamma^* = \mathbf{V}\mathbf{D}_\mu^{\frac{1}{2}}$, estas novas coordenadas correspondem ao método gráfico *Biplot* (Gabriel, 1971), mas não trabalharemos com estas coordenadas, já que nosso interesse está voltado para as propriedades geométricas da representação discutida nesta seção.

2.4.2 Auxílios à Interpretação

Nem sempre é possível graficar a imagem completa, usualmente, tentamos imaginar a configuração completa dos pontos, a partir dos gráficos de uma, duas, ou no máximo três dimensões. Para interpretar uma imagem parcial ("projetada sobre os primeiros eixos"), é necessário conhecer quão bem representados estão os pontos, e também a significação dos eixos sobre os quais temos projetado a nuvem. Começaremos então, por entender como a inércia é decomposta, que é a base para a qualidade de representação e a significação dos eixos.

A Decomposição da Inércia

Como consequência de (13), (14) e (53), a inércia total de cada nuvem de pontos, pode ser decomposta ao longo dos eixos principais, ou ao longo dos vetores que representam os pontos, de forma similar e simétrica. Isto dá a decomposição da inércia, análoga à decomposição da variância para cada nuvem de pontos, como ilustram as tabelas 1 e 2. Estas tabelas formam o suporte numérico para o gráfico. Cada coluna das tabelas, representa a decomposição da inércia contabilizada pelo eixo correspondente, e cada linha destas tabelas, representa a decomposição da inércia dada por cada categoria linha ou coluna. Destas tabelas, e de (17), temos:

$$\text{tr}(\mathbf{D}_\mu^2) = In = \chi^2/n \quad (63)$$

Tabela 1: Decomposição da Inércia da nuvem de linhas.

Linhas	Eixos				Total
	1	2	...	D	
1	$p_{1.} f_{11}^2$	$p_{1.} f_{12}^2$	...	$p_{1.} f_{1D}^2$	$p_{1.} \sum_a f_{1a}^2$
2	$p_{2.} f_{21}^2$	$p_{2.} f_{22}^2$	...	$p_{2.} f_{2D}^2$	$p_{2.} \sum_a f_{2a}^2$
...					
I	$p_{I.} f_{I1}^2$	$p_{I.} f_{I2}^2$	...	$p_{I.} f_{ID}^2$	$p_{I.} \sum_a f_{Ia}^2$
Total	μ_1^2	μ_2^2	...	μ_D^2	In

Tabela 2: Decomposição da Inércia da nuvem de colunas.

Colunas	Eixos				Total
	1	2	...	D	
1	$p_{.1} g_{11}^2$	$p_{.1} g_{12}^2$	...	$p_{.1} g_{1D}^2$	$p_{.1} \sum_a g_{1a}^2$
2	$p_{.2} g_{21}^2$	$p_{.2} g_{22}^2$	...	$p_{.2} g_{2D}^2$	$p_{.2} \sum_a g_{2a}^2$
...					
J	$p_{.J} g_{J1}^2$	$p_{.J} g_{J2}^2$	...	$p_{.J} g_{JD}^2$	$p_{.J} \sum_a g_{Ja}^2$
Total	μ_1^2	μ_2^2	...	μ_D^2	In

A expressão (63) mostra que a AC decompõe a estatística χ^2 para testar independência na matriz. De fato, a contribuição das parcelas, indica em qual direção se dão os desvios da independência. Naturalmente, procederemos a calcular os coeficientes que nos auxiliarão na interpretação. Para cada elemento linha e coluna, são calculados:

A Contribuição Absoluta

Expressa a participação que tem o elemento na inércia de um eixo.

- A *Contribuição Absoluta* da i -ésima linha ao α -ésimo eixo, está dada por:

$$\begin{aligned} ca_{[R]ia} &= P_{i.} f_{ia}^2 / \mu_\alpha^2 \\ &= P_{i.} \phi_{ia}^2, \end{aligned} \quad (64)$$

onde $\mu_\alpha^2 = \sum_i P_{i.} f_{ia}^2$ ($\alpha = 1, \dots, D$) é a inércia explicada pelo eixo α . Observemos que, para todo $\alpha = 1, \dots, D$: $\sum_i ca_{[R]ia} = 1$.

- A *Contribuição Absoluta* da j -ésima coluna ao α -ésimo eixo, está dada por:

$$\begin{aligned} ca_{[C]ja} &= P_{.j} g_{ja}^2 / \mu_\alpha^2 \\ &= P_{.j} \Gamma_{ja}^2, \end{aligned} \quad (65)$$

onde $\mu_\alpha^2 = \sum_j P_{.j} g_{ja}^2$ ($\alpha = 1, \dots, D$) é a inércia explicada pelo eixo α . Observemos que, para todo $\alpha = 1, \dots, D$: $\sum_j ca_{[C]ja} = 1$.

Usa-se o termo *absoluta*, devido que a contribuição está afetada pela massa.

A Contribuição Relativa

Expressa a participação de um eixo na explicação de um elemento. Mede a qualidade de representação do elemento pelo eixo.

- A *Contribuição Relativa* do α -ésimo eixo à i -ésima linha, está dada por:

$$\begin{aligned} cr_{[R]ia} &= f_{ia}^2 / \sum_\alpha f_{ia}^2 \\ &= \phi_{ia}^2 / \sum_\alpha \phi_{ia}^2, \end{aligned} \quad (66)$$

observemos que, esta medida é equivalente a tomar o quadrado do cosseno do ângulo entre o ponto em coordenadas principais e o eixo principal, ou o quadrado do cosseno do ângulo entre o

ponto em coordenadas estandardizadas e o fator, e cumpre para todo $i = 1, \dots, I$: $\sum_{\alpha} c_{r|\mathbf{r}|i\alpha} = 1$.

- A *Contribuição Relativa* do α -ésimo eixo à j -ésima coluna, está dada por:

$$\begin{aligned} c_{r|\mathbf{c}|ja} &= g_{ja}^2 / \sum_{\alpha} g_{ja}^2 \\ &= \Gamma_{ja}^2 / \sum_{\alpha} \Gamma_{ja}^2, \end{aligned} \quad (67)$$

observemos que, esta medida é equivalente a tomar o quadrado do cosseno ângulo entre o ponto e o eixo ou o fator, e cumpre para todo $j = 1, \dots, J$: $\sum_{\alpha} c_{r|\mathbf{c}|ja} = 1$.

A Qualidade de Representação

É a participação de um sub-conjunto de eixos na explicação de um elemento.

- A *Qualidade de Representação* de uma linha i por um sub-conjunto de eixos $\mathbf{S}$, está dada por:

$$q_{r|\mathbf{r}|i\mathbf{S}} = \sum_{\alpha \in \mathbf{S}} c_{r|\mathbf{r}|i\alpha}. \quad (68)$$

- A *Qualidade de Representação* de uma coluna j por um sub-conjunto de eixos $\mathbf{S}$, está dada por:

$$q_{r|\mathbf{c}|j\mathbf{S}} = \sum_{\alpha \in \mathbf{S}} c_{r|\mathbf{c}|j\alpha}. \quad (69)$$

2.4.3 Interpretação dos Eixos

A interpretação dos eixos é fundamental, a visão direta do gráfico pode induzir a erros. Para evitar erros, será necessário tomar em conta as contribuições absolutas e relativas. A seguir, sugeriremos uma série de passos que poderão nos conduzir a uma interpretação correta:

- a. Em primeiro lugar, se buscam aqueles pontos de maior ca .

- b. Entre eles, se separam os pontos que se projetam do lado positivo dos que intervêm do lado negativo, que estarão em oposição.
- c. Estuda-se a qualidade de representação *cr* destes pontos. Se um ponto tem uma *cr* pequena, é de se supor que tenha um papel importante sobre outro eixo, para seu estudo, será necessário considerar o conjunto dos eixos.
- d. Buscam-se aqueles pontos que estejam bem representados (i.e. têm um *cr* alto).

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 um significado aos eixos. É relativamente fácil, quando há elementos muito bem representados sobre os eixos e opostos entre si. Frequentemente, se interpreta por oposição de adjetivos, como menção A contra menção E, ou mercado rural contra mercado urbano, ou seja, através de contrastes entre categorias.

2.4.4 Os Elementos Suplementares

A interpretação da análise pode ser facilitada ou enriquecida com a representação gráfica dos *elementos suplementares* ou *ilustrativos*. São elementos que não intervêm na análise, não participam na formação dos eixos, mas podem ser projetados sobre os gráficos obtidos, de modo que sua localização seja um auxílio à interpretação.

- A projeção de uma linha suplementar s , de valores n_{sj} se obtém das fórmulas baricentricas:

$$f_{sa} = \mathbf{R}_s' \mathbf{G} \mathbf{D}_\mu^{-1} , \quad (70)$$

onde $\mathbf{R}_s = 1/n_{s.} (n_{s1} , \dots , n_{sj})'$, com $n_{s.} = \sum_j n_{sj}$.

- Análogamente, a projeção de uma coluna suplementar s , de valores n_{is} , está dada por:

$$E_{s\alpha} = \mathbf{c}_s' \mathbf{F} \mathbf{D}_\mu^{-1}, \quad (71)$$

onde $\mathbf{c}_s = 1/n_{.s} (n_{1s}, \dots, n_{Is})'$, com $n_{.s} = \sum_i n_{is}$.

O uso dos elementos suplementares é muito útil e inclusive, em alguns casos, necessário. Geralmente, são utilizados para:

- Representar centros de gravidade de algumas classes, em especial quando o número de linhas é excessivamente grande.
- Para representar a caracterização de distintos grupos de indivíduos (um grupo pode ser definido pelo cruzamento de uma linha com uma coluna).
- Para representar um elemento que perturbava uma análise anterior (por ser muito diferente ou ter um grande peso) impedindo a interpretação clara da análise.
- Para introduzir novas informações exógenas que facilitam a interpretação dos fatores.

A interpretação não é uma mera técnica, se não, também leva consigo algo de arte. Geralmente, uma vez realizada uma análise, se observa a conveniência de outras novas, nas quais pode ser que, consideremos alguma(s) categoria(s) como suplementar(es), ou incluamos outras como efectivas.

Na seções 4.2 e 4.3 temos colocado exemplo simples de aplicação da AC.

2.5 Análise de Correspondência Generalizada

A AC de uma tabela pode ser vista como a análise da diferença entre a tabela e o modelo de independência. Existem casos em que o estudo do problema se adapta melhor a algum modelo diferente do modelo de independência e, portanto, será preferível analisar os resíduos de tal modelo. Para lograrmos isto, faremos uso de uma generalização da análise de correspondência proposta por Escofier (1983, 1984), que exporemos a seguir:

2.5.1 Os Dados

Dada uma tabela de correspondência $\mathbf{P}$, definiremos a matriz arbitrária $\mathbf{Q}$, da mesma ordem que $\mathbf{P}$:

$$\mathbf{Q} = [q_{ij}]_{I \times J}, \quad (72)$$

como a tabela de um modelo para as observações em $\mathbf{P}$, tal que $\mathbf{1}' \mathbf{Q} \mathbf{1} = 1$. Este modelo, pode ser qualquer um, como o modelo de independência, o modelo de quase-independência, o modelo de simetria, ou pode ser construída a partir de outras variáveis mediante de uma regressão, etc...

Daqui definiremos os vetores $\mathbf{r}_Q$ e $\mathbf{c}_Q$ de $\mathbb{R}^I$ e $\mathbb{R}^J$ respectivamente como as marginais de $\mathbf{Q}$. i.e.:

$$\begin{aligned} \mathbf{r}_Q &= (q_{1.}, q_{2.}, \dots, q_{I.})' = \mathbf{Q} \mathbf{1}_J \\ \mathbf{c}_Q &= (q_{.1}, q_{.2}, \dots, q_{.J})' = \mathbf{Q}' \mathbf{1}_I. \end{aligned} \quad (73)$$

observemos que, $\mathbf{1}_I' \mathbf{r}_Q = 1$ e $\mathbf{1}_J' \mathbf{c}_Q = 1$.

Definiremos também os vetores $\mathbf{r}_g$ e $\mathbf{c}_g$ de $\mathbb{R}^I$ e $\mathbb{R}^J$ respectivamente, como vetores arbitrários com coordenadas maiores que zero e tais que:

$$\begin{aligned}\mathbf{1}_I' \mathbf{r}_g &= 1 \\ \mathbf{1}_J' \mathbf{c}_g &= 1.\end{aligned}\tag{74}$$

Estes podem ser uma combinação das marginais de $\mathbf{P}$ e $\mathbf{Q}$, etc.... Para estes vetores, definiremos as matrizes diagonais $\mathbf{D}_{rg} = \text{diag}\{\mathbf{r}_g\}$ e $\mathbf{D}_{cg} = \text{diag}\{\mathbf{c}_g\}$.

Formação dos Pontos no Espaço

Agora estamos interessados em representar a diferença entre a tabela e o modelo.

- Em $\mathbb{R}^I$ formaremos uma nuvem de I pontos que representarão a diferença entre as linhas de $\mathbf{P}$ e $\mathbf{Q}$:

$$\mathbf{R}_g = \mathbf{D}_{rg}^{-1} (\mathbf{P} - \mathbf{Q})\tag{75}$$

i.e. $\mathbf{R}_g = [\mathbf{r}_{g1} \mid \mathbf{r}_{g2} \mid \dots \mid \mathbf{r}_{gI}]'$, onde para todo i ($i = 1, \dots, I$) o perfil linha generalizado:

$$\mathbf{r}_{gi} = \left[\frac{p_{i1} - q_{i1}}{r_{gi}}, \frac{p_{i2} - q_{i2}}{r_{gi}}, \dots, \frac{p_{iJ} - q_{iJ}}{r_{gi}} \right]'$$

- Em $\mathbb{R}^J$ formaremos uma nuvem de J pontos que representarão a diferença entre as colunas de $\mathbf{P}$ e $\mathbf{Q}$:

$$\mathbf{C}_g = \mathbf{D}_{cg}^{-1} (\mathbf{P} - \mathbf{Q})'\tag{76}$$

i.e. $\mathbf{C}_g = [\mathbf{c}_{g1} \mid \mathbf{c}_{g2} \mid \dots \mid \mathbf{c}_{gJ}]'$, onde para todo j ($j = 1, \dots, J$) o perfil coluna generalizado:

$$\mathbf{c}_{gj} = \left[\frac{p_{1j} - q_{1j}}{c_{gj}}, \frac{p_{2j} - q_{2j}}{c_{gj}}, \dots, \frac{p_{Ij} - q_{Ij}}{c_{gj}} \right]'$$

Observemos que, se $\mathbf{r}_g = \mathbf{r}_0 = \mathbf{r}$ e $\mathbf{c}_g = \mathbf{c}_0 = \mathbf{c}$, então estes representam a diferença entre os perfis da tabela e do modelo.

As Massas

Em $\mathbb{R}^J$, atribuiremos a cada ponto i ($i = 1, \dots, I$), uma massa igual a r_{gi} , assim, vetor das massas das linhas é $\mathbf{r}_g$. Análogamente, em $\mathbb{R}^I$, atribuiremos a cada ponto j ($j = 1, \dots, J$), uma massa igual a c_{gj} , assim, o vetor das massas das colunas é $\mathbf{c}_g$.

A Distância

Usaremos a distância χ^2 , em $\mathbb{R}^J$, a métrica euclideana ponderada por $\mathbf{D}_{\mathbf{c}_g}^{-1}$, e em $\mathbb{R}^I$, métrica euclideana ponderada por $\mathbf{D}_{\mathbf{r}_g}^{-1}$. Assim, a distância entre dois pontos linha i e l ($i, l = 1, \dots, I$) e a distância entre dois pontos coluna j e k ($j, k = 1, \dots, J$), são respectivamente:

$$\begin{aligned} d^2_{\{\mathbf{r}_g\}}(i, l) &= (\mathbf{r}_{gi} - \mathbf{r}_{gl})' \mathbf{D}_{\mathbf{c}_g}^{-1} (\mathbf{r}_{gi} - \mathbf{r}_{gl}) \\ d^2_{\{\mathbf{c}_g\}}(j, k) &= (\mathbf{c}_{gi} - \mathbf{c}_{gl})' \mathbf{D}_{\mathbf{r}_g}^{-1} (\mathbf{c}_{gi} - \mathbf{c}_{gl}) \end{aligned} \quad (77)$$

Equivalência Distribucional

Nesta definição, da distância, não é difícil demonstrar que a propriedade de equivalência distribucional, ainda se cumpre:

- Se duas linhas de $\mathbf{R}_g$ são iguais, elas podem ser substituídas por uma só, a qual é igual a elas, mas com uma massa igual a soma das massas, sem que isto altere as distâncias entre as linhas de $\mathbf{C}_g$.
- Se duas linhas de $\mathbf{C}_g$ são iguais, elas podem ser substituídas por uma só, a qual é igual a elas, mas com uma massa igual a

soma das massas, sem que isto altere as distâncias entre as linhas de $\mathbf{R}_g$.

Assim, se duas categorias "iguais" são reunidas em uma só, os resultados em geral não serão modificados.

Os Centróides

Os centróides dos pontos linha e coluna estão dados por:

$$\begin{aligned}\text{Centróide}_{\{R_g\}} &= \mathbf{R}_g' \mathbf{r}_g = (\mathbf{P} - \mathbf{Q})' \mathbf{1} = \mathbf{c} - \mathbf{c}_Q, \\ \text{Centróide}_{\{C_g\}} &= \mathbf{C}_g' \mathbf{c}_g = (\mathbf{P} - \mathbf{Q}) \mathbf{1} = \mathbf{r} - \mathbf{r}_Q.\end{aligned}\tag{78}$$

Observemos que, se as marginais da tabela e do modelo são iguais, então, as nuvens estarão centradas na origem. Mas, se estivermos no caso do centróide não coincidir com a origem, a nuvem de pontos não será bem representada pelas projeções sobre os primeiros eixos, a menos que as marginais do modelo não sejam muito diferentes das da tabela observada.

2.5.2 A Análise

As coordenadas principais são, por definição, as projeções das nuvens sobre os eixos de inércia. Tais eixos são calculados da origem, que nem sempre contém o centro de gravidade da nuvem. Assim, as coordenadas principais são:

$$\mathbf{F} = \mathbf{R}_g \mathbf{D}_{cg}^{-\frac{1}{2}} \mathbf{V} \quad \text{e} \quad \mathbf{G} = \mathbf{C}_g \mathbf{D}_{cg}^{-\frac{1}{2}} \mathbf{U}.\tag{79}$$

onde:

$$\mathbf{F} = [\mathbf{F}_{(1)} \mid \mathbf{F}_{(2)} \mid \dots \mid \mathbf{F}_{(D)}],$$

$$\mathbf{G} = [\mathbf{G}_{(1)} \mid \mathbf{G}_{(2)} \mid \dots \mid \mathbf{G}_{(D)}],$$

$$\mathbf{V} = [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_D], \text{ com } \mathbf{V}'\mathbf{V} = \mathbf{I}_J \text{ e}$$

$$\mathbf{U} = [\mathbf{u}_1 \mid \dots \mid \mathbf{u}_D], \text{ com } \mathbf{U}'\mathbf{U} = \mathbf{I}_I,$$

$$\text{para } D \leq \min \{i, j\}$$

e onde, as matrizes $\mathbf{U}$ e $\mathbf{V}$, são obtidas da decomposição da matriz $\mathbf{Y} = \mathbf{D}_{rk}^{-\frac{1}{2}} (\mathbf{P} - \mathbf{Q}) \mathbf{D}_{ck}^{-\frac{1}{2}}$, segundo:

$$\mathbf{Y} = \mathbf{U} \mathbf{D}_{\mu} \mathbf{V}', \quad (80)$$

com $\mathbf{D}_{\mu} = \text{diag}\{\mu\}$, com $\mu = (\mu_1, \dots, \mu_D)'$
e $\mu_1 \geq \dots \geq \mu_D > 0$.

Observe que, neste caso não necessariamente perdemos uma dimensão como na AC.

As Equações de Transição

De forma análoga, que na AC, temos as equações de transição:

$$\begin{aligned} \mathbf{F} &= \mathbf{R}_r \mathbf{G} \mathbf{D}_{\mu}^{-1} & \text{e} & \quad \mathbf{G} = \mathbf{C}_r \mathbf{F} \mathbf{D}_{\mu}^{-1} \text{ ou} \\ \Phi &= \mathbf{R}_r \mathbf{\Gamma} \mathbf{D}_{\mu}^{-1} & \text{e} & \quad \mathbf{\Gamma} = \mathbf{C}_r \Phi \mathbf{D}_{\mu}^{-1}. \end{aligned} \quad (81)$$

onde, $\mathbf{F}$ e $\mathbf{G}$ são as coordenadas principais e Φ e $\mathbf{\Gamma}$ são as coordenadas standardizadas nos espaços $\mathbb{R}^I$ e $\mathbb{R}^J$ respectivamente.

A partir destas, podemos interpretar as posições relativas entre os pontos linha e coluna, da seguinte maneira:

- Um ponto linha, estará situado ao lado daqueles pontos coluna aos que se associa mais na tabela do que no modelo ($p_{ij} > q_{ij}$) e ao lado oposto daqueles aos que se associa menos na tabela do que no modelo ($p_{ij} < q_{ij}$).
- Um ponto coluna, estará situado ao lado daqueles pontos linha aos que se associa mais na tabela do que no modelo ($p_{ij} > q_{ij}$) e ao lado oposto daqueles aos que se associa menos na tabela do que no modelo ($p_{ij} < q_{ij}$).

Reconstrução da Tabela

É mais fácil discutir esta generalização observando sua fórmula de reconstrução, que é análoga, com as coordenadas principais a (47) e com as coordenadas estandardizadas a (60):

$$\begin{aligned} \mathbf{P} &= \mathbf{Q} + \mathbf{D}_{rk} \mathbf{F} \mathbf{D}_{\mu}^{-1} \mathbf{G}' \mathbf{D}_{cg} \text{ , ou} \\ \mathbf{P} &= \mathbf{Q} + \mathbf{D}_{rk} \mathbf{\Phi} \mathbf{D}_{\mu} \mathbf{\Gamma}' \mathbf{D}_{cg} \text{ .} \end{aligned} \tag{82}$$

A comparação das equações em (82) com as equações em (47) e (60), revela que a AC está generalizada em dois sentidos: em primeiro lugar, é possível decompor a diferença entre a tabela $\mathbf{P}$ e um modelo $\mathbf{Q}$, diferente do modelo de independência $\mathbf{E} = \mathbf{rc}'$; em segundo lugar, os pesos $\mathbf{r}_g$ e $\mathbf{c}_g$ dos pontos linha e os pontos coluna, não são necessariamente definidos pelas marginais de $\mathbf{P}$.

Escoufier (1985) discute também outras alternativas para a generalização dada na equação (82), mas só usaremos esta generalização particular, porque ela é conveniente para a decomposição de resíduos de modelos log-lineares.

2.5.3 Casos Particulares

Se $\mathbf{Q}$ corresponde ao modelo de independência (i.e. $q_{i,j} = p_{i.} p_{.j}$), as marginais do modelo são iguais às marginais da tabela (i.e.: $\mathbf{r}_0 = \mathbf{r}$ e $\mathbf{c}_0 = \mathbf{c}$). Se além disso, $\mathbf{r}_g = \mathbf{r}$ e $\mathbf{c}_g = \mathbf{c}$, então estaremos trabalhando com a análise de correspondência usual.

Voltando ao caso mais geral, se somente exigimos que, $\mathbf{r}_g = \mathbf{r}_0 = \mathbf{r}$ e $\mathbf{c}_g = \mathbf{c}_0 = \mathbf{c}$, os resultados podem ser obtidos pela análise de correspondência ordinária aplicada à tabela:

$$\mathbf{P} = \mathbf{Q} + \mathbf{rc}' \text{ .}$$

Com efeito, as nuvens associadas à AC de tal tabela, são as mesmas nuvens definidas em (75) e (76). Elas representam as diferenças entre os perfis da tabela e do modelo.

Veja os exemplos na seções 4.2 e 4.3 que, sob diferentes circunstâncias, a aplicação da ACG para os distintos modelos é útil.

3. Relacionando a Análise de Correspondência com Modelos Log-Lineares

3.1 Introdução

A maior diferença entre a *Análise de Correspondência* (AC) e a maioria das outras técnicas para dados categóricos, está em que a AC recupera a informação dos dados de forma gráfica, obtendo-se a decomposição da estrutura dos dados em forma ótima e sem uso de modelos probabilísticos: A matriz dos dados é aproximada por uma matriz de posto menor, usando o método da soma de quadrados mínimos. Tem portanto, caráter exploratório e, se aplica a estudos tanto experimentais, quanto observacionais. A *Análise Log-linear* (ALL) é uma técnica basicamente de caráter confirmatório: por exemplo, se assume uma distribuição sob a qual os dados são coletados, logo se assume um modelo hipotético, as estimativas são feitas sob a suposição de que o modelo é verdadeiro, e finalmente, estas estimativas são comparadas com as frequências observadas para avaliar o modelo.

Veremos aqui que, podemos integrar AC e ALL para a análise de tabelas de contingência. Uma primeira abordagem para comparar a AC com outras abordagens que se baseiam em modelos foi feita por Goodman (1979a, 1981a,b), ele tratou sobre quais condições os resultados da AC são similares aos do

modelo. Logo depois, a brecha entre AC e as abordagens baseadas em modelos, foi estreitada por Goodman (1985a,b, 1986) e Gilula e Haberman (1986), que discutiram as relações entre a AC e os modelos de associação RC (modelo log-multiplicativo linha-coluna RC), e introduziram uma versão da AC como um modelo. Recentemente, van der Heijden e de Leeuw (1985), van der Heijden (1986), Worsley (1987), van der Heijden e Worsley (1988), van der Heijden, de Falguerolles e de Leeuw (1989), Choulakian (1988), e Hudon (1990), entre outros, fizeram uso destes resultados para combinar AC e ALL na análise de tabelas de contingência.

Nesta seção, faremos uma breve referência a ALL, e mostraremos como a AC e a ALL podem relacionar-se. Indicaremos também, como poderemos analisar, combinando ambos os métodos. Nesta abordagem, a ALL é usada para saber a importância dos efeitos de interação de um modelo, e logo a AC é usada para estudar os resíduos do modelo.

3.2 Análise Log-linear

A *Análise Log-linear* (ALL) é um método bastante conhecido para o estudo de relações estruturais entre variáveis em uma tabela de contingência $\mathbf{P}_{I \times J}$. Aqui, introduziremos somente a notação para o caso de duas ou três variáveis, já que com mais variáveis a generalização seria feita de forma natural.

Em geral, com uma matriz de duas entradas, um modelo log-linear tem a forma:

$$\log \pi_{ij} = u + u_{1(i)} + u_{2(j)} + u_{12(i,j)} \quad (83)$$

onde π_{ij} denota a probabilidade para a casela (i,j) ($i = 1, \dots, I$, $j = 1, \dots, J$), e u , $u_{1(i)}$, $u_{2(j)}$ e $u_{12(i,j)}$ são os parâmetros do modelo. Existem várias alternativas para o modelo acima, de modo que os parâmetros devem ser restringidos para identificar o modelo (veja Bishop et al. (1975)).

Para denotar os modelos log-lineares, seguiremos a convenção usual. A seguir daremos alguns exemplos de modelos log-lineares, em especial para

alguns tipos de independência, seguidos da equivalência em termos de probabilidade e a notação usual, que é mais compacta e será utilizada adiante:

a. Modelo Saturado:

$$\log \pi_{ij} = u + u_{1(i)} + u_{2(j)} + u_{12(ij)}$$

$$\text{equivale a: } \pi_{ij} = p_{ij}$$

notação: [12].

b. Modelo de Independência:

$$\log \pi_{ij} = u + u_{1(i)} + u_{2(j)}$$

$$\text{equivale a: } \pi_{ij} = \pi_{i.} \pi_{.j}$$

notação: [1][2].

Os seguintes exemplos, para tabelas de três entradas:

a. Modelo de Independência Mutua:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)}$$

$$\text{equivale a: } \pi_{ijk} = \pi_{i..} \pi_{.j.} \pi_{..k}$$

notação: [1][2][3].

b. Modelo de Independência Conjunta:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{13(ik)}$$

$$\text{equivale a: } \pi_{ijk} = \pi_{.j.} \pi_{i.k}$$

notação: [2][13].

c. Modelo de Independência Condicional:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{13(ik)} + u_{23(jk)}$$

$$\text{equivale a: } \pi_{ij/k} = \pi_{i./k} \pi_{.j/k}$$

notação: [13][23].

d. Modelo de Marginais Independentes:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)}$$

$$\text{equivale a: } \pi_{ij.} = \pi_{i..} \pi_{.j.}$$

notação: [1][2].

A interpretação dos parâmetros u , é às vezes difícil, especialmente se eles são muitos ou, quando representam interações de alta ordem que não podem ser desprezadas. Usualmente, é de grande ajuda tentar interpretar por meio da representação gráfica dos parâmetros ou, usando a razão de odds. Outra maneira de abordar o problema de interpretar um grande número de parâmetros log-lineares, é de restringir, de alguma forma, os parâmetros de interação, por exemplo, que tenham forma produto. Isto é feito no modelo de

associação RC (Goodman, 1979a, b, 1981a, b, 1985a, b, 1986, 1991; Andersen, 1980). De modo que, quando o número de categorias é grande, o número de parâmetros a interpretar pode ser consideravelmente reduzido. Na seguinte seção, apresentaremos uma alternativa, relacionada a esta última, para a interpretação.

3.3 Relações entre a Análise de Correspondência e alguns Modelos Log-lineares

A representação da tabela $\mathbf{P}$ com a AC em d dimensões ($d = 1, \dots, D$, onde $D \leq \min\{I, J\} - 1$), dada na expressão (48), pode ser reescrita como:

$$\mathbf{P}^{(d)} = \mathbf{D}_r (\mathbf{1}_I \mathbf{1}_J' + \Phi^{(d)} \mathbf{D}_\mu^{(d)} \Gamma^{(d)}) \mathbf{D}_c, \quad (84)$$

isto é:

$$p^{(d)}_{ij} = p_{i.} p_{.j} (1 + \sum_a^d \mu_a \phi_{ia} \Gamma_{ja}),$$

onde $p^{(d)}_{ij}$ é a reconstrução de ordem d da casela (i, j) .

Esta equação pode ser reescrita tomando o log de $p^{(d)}_{ij}$, para tomar a forma do modelo de associação RC, que é uma classe de modelos log-lineares (veja Goodman, 1979a, b, 1981a, b, 1985a, b, 1986, 1991; Andersen, 1980), i.e.:

$$\log(p^{(d)}_{ij}) = u + u_{1(i)} + u_{2(j)} + u_{12(ij)}, \quad (85)$$

onde:

$$u = 0,$$

$$u_{1(i)} = \log p_{i.},$$

$$u_{2(j)} = \log p_{.j} \text{ e}$$

$$u_{12(ij)} = \log(1 + x) \quad , \text{ com } x = \sum_a^d \mu_a \phi_{ia} \Gamma_{ja}.$$

Observações

A partir dos resultados acima, vejamos algumas observações para começar a entender de que maneira a AC e o modelo de associação RC estão relacionados.

- a. Notemos que $u_{12(ij)}$ é uma função crescente relativamente a x .
- b. Em geral, na representação simultânea da AC, se os pontos linha i e coluna j estiverem situados em lados opostos da nuvem, o valor de x será negativo, e por conseguinte $u_{12(ij)}$ também. Reciprocamente, se os pontos i e j estiverem situados do mesmo lado da nuvem, os valores x e $u_{12(ij)}$ serão positivos. E, se um dos pontos estiver bem perto da origem, ou se os vetores que os representam fossem ortogonais entre si, a interação devida a este ponto com qualquer outro ponto da outra nuvem será praticamente nula.
- c. Devido a que, as coordenadas de Φ e Γ estão centradas na origem e normalizadas segundo a equação (57) (i.e. para cada α : $\sum_i p_i \phi_{i\alpha}^2 = 1 = \sum_j p_{.j} \Gamma_{j\alpha}^2$), o tamanho de x depende principalmente dos valores de $\mu^{(d)}$, que indicam o afastamento da independência (veja expressão (63)). Quanto mais distante estiverem os dados da independência, μ_α ($\alpha = 1, \dots, d$) estará mais distante de zero, e por conseguinte, a interação $u_{12(ij)}$ será não desprezível.

Agora, com as observações acima, podemos ter fortes indícios, sem ainda testar, de quais são as interações importantes e suas devidas interpretações, se existirem.

Casos Particulares

- Escoufier (1982), notou que, se x ($x = \sum_{\alpha}^d \mu_{\alpha} \phi_{i\alpha} \Gamma_{j\alpha}$) for pequeno em comparação a unidade (de forma que $\log(1 + x) \approx x$), então:

$$\log(p^{(d)}_{ij}) \approx u_{1(i)} + u_{2(j)} + x. \quad (86)$$

Este modelo é chamado de modelo log-bilinear (veja Goodman, 1986). A condição de Escoufier é satisfeita quando o desvio da independência não é muito grande i.e., quando a soma das inércias é pequena com relação a um.

- Se $d = 1$, então não tem importância o tamanho de μ_1 , a equação (86) estará intimamente relacionada ao modelo de associação RC (onde os parâmetros de interação têm forma produto), que foi introduzido por Goodman (1981a, 1985a), e se expressa assim:

$$\log(\pi_{ij}) = u + u_{1(i)} + u_{2(j)} + \phi u_{1(i)}^* u_{2(j)}^* \quad (87)$$

onde, $\phi \approx \mu_1$, $u_{1(i)}^* \approx \phi_{i1}$, e $u_{2(j)}^* \approx \Gamma_{j1}$, e:

$$\sum_i p_{i.} u_{1(i)}^{*2} = 1 = \sum_j p_{.j} u_{2(j)}^{*2}.$$

A diferença entre as equações (86) e (87), é que a equação (86) é uma aproximação com d fatores, enquanto que a equação (87) usa somente o primeiro fator. Comparado com o modelo (83), o modelo (87) pode ser interpretado como um modelo que admite interação log-multiplicativa, onde:

$$u_{12(ij)} = \phi u_{1(i)}^* u_{2(j)}^*.$$

Um modelo de associação RC multifatorial, foi introduzido por Goodman (1985, 1986). Para tabelas de maior dimensão, Clogg (1982), de Leeuw (1983), Agresti (1984) e Pannekoek (1985), introduziram modelos log-lineares com restrições nas interações.

Daqui, podemos concluir que, na AC a interação está decomposta na forma log-multiplicativa: os gráficos, resultados da AC, mostram aproximações dos parâmetros log-multiplicativos. Quanto maior for a interação positiva entre uma categoria linha e uma categoria coluna, os pontos que os representam se localizarão na periferia da nuvem e perto um do outro, reciprocamente, quanto maior for o módulo da interação negativa entre uma categoria linha e uma categoria coluna, mais longe (em lados opostos do gráfico) estarão os pontos que os representam. Isto mostra que AC pode dar indicações sobre os parâmetros de interação log-linear (veja Worsley (1987), van der Heijden (1987) e van der Heijden e Worsley (1988)).

3.4 Análise de Correspondência como uma Técnica para Decompor Resíduos de alguns Modelos de Independência Log-linear para Tabelas Multi-Entrada

Na secção 2.3.4, foi enfatizado que AC decompõe os resíduos do modelo de independência, de modo que, para uma tabela de duas entradas com proporções p_{ij} , está claro que a AC pode ser interpretada em termos da diferença entre os modelos [12] e [1][2]. Recentemente, foi mostrado que AC também está relacionada a ALL no caso de tabelas de contingência multi-entrada (van der Heijden e de Leeuw, 1985).

Os resultados são muito interessantes para a análise de uma tabela multi-entrada: transformando a tabela em uma tabela múltipla, i.e. uma tabela de duas entradas construída concatenando os pedacos da tabela multi-entrada.

Por exemplo, consideremos uma tabela $\mathbf{P}$ de dimensões $I \times J \times K$ (três entradas), com proporções p_{ijk} ($i = 1, \dots, I$, $j = 1, \dots, J$, $k = 1, \dots, K$). A tabela múltipla de duas entradas $\mathbf{P}^{i[jk]}$ é construída codificando as variáveis 2 e 3 interactivamente. Denotaremos por $\mathbf{E}^{i[jk]}$ à esperança de $\mathbf{P}^{i[jk]}$ sob hipótese de independência da primeira variável contra a conjunta da segunda com a terceira. Denotaremos as caselas de $\mathbf{P}^{i[jk]}$ e $\mathbf{E}^{i[jk]}$ como $p_{i[jk]}$ e $e_{i[jk]}$ respectivamente. Os valores $e_{i[jk]}$ são iguais às estimativas de máxima verossimilhança da esperança sob o modelo [1][23]:

$$\hat{\pi}_{ijk} = p_{i..} p_{.jk} = p_{i[.]} p_{.[jk]} = e_{i[jk]}$$

Assim, quando a matriz $\mathbf{P}$, de 3 entradas, é convertida na tabela múltipla $\mathbf{P}^{i[jk]}$, a AC pode ser interpretada em termos da diferença entre os modelos log-lineares [123] e [1][23]. Análogamente, quando as tabelas múltiplas $\mathbf{P}^{i[ik]}$ ou $\mathbf{P}^{k[ij]}$, forem construídas, os valores em $\mathbf{E}$ seguirão os modelos [2][13] ou [3][12] respectivamente.

Em geral, quando temos uma tabela multi-entradas, podemos construir uma tabela múltipla $\mathbf{P}^{[a][b]}$, onde a e b indexam os elementos de dois grupos de variáveis $\mathbf{A}$ e $\mathbf{B}$ codificados interactivamente. $\mathbf{A}$ e $\mathbf{B}$ constituem a tabela multi-entrada onde $\mathbf{A} \cap \mathbf{B} = \emptyset$. Neste caso, a AC pode ser interpretada em termos da diferença entre os modelos $[\mathbf{AB}]$ e $[\mathbf{A}][\mathbf{B}]$: os elementos de $\mathbf{E}$ são iguais a $e_{ab} = p_{a..} p_{.b} = \hat{\pi}_{ab}$.

Portanto, a AC simples, aplicada a uma tabela múltipla, pode ser interpretada como uma ferramenta para analisar resíduos do modelo log-linear de independência, que corresponde a construção da tabela múltipla, a partir da tabela multi-entrada. Não obstante, em algumas aplicações, estes modelos não serão aqueles nos quais estamos mais interessados. Outros modelos, permitem só independência condicional ou ausência de interações de alta ordem.

Com a generalização da AC discutida na seção 2.5, é possível analisar a diferença entre modelos log-lineares menos restritos e o modelo saturado. As proporções estimadas $\hat{n}_{ijk}$ são tomadas como os elementos de $\mathbf{Q}$ (veja a equação (82)), e as marginais da tabela múltipla de $\mathbf{P}$ (que são iguais às de $\mathbf{Q}$) são tomadas como os pesos, assim: $\mathbf{r}_g = \mathbf{r}_0 = \mathbf{r}$ e $\mathbf{c}_g = \mathbf{c}_0 = \mathbf{c}$. Como foi observado na seção 2.5, nesta aplicação, as soluções da AC podem ser obtidas utilizando qualquer programa para AC simples aplicada à tabela: $\mathbf{P} - \mathbf{Q} + \mathbf{r}\mathbf{c}'$. Veja os exemplos nas seções 4.2.1 e 4.3.2.

O uso de modelos log-lineares menos restritos têm implicações geométricas interessantes (veja Escofier (1983) e van der Heijden (1987)). Consideremos a análise da tabela múltipla $\mathbf{P}^{[i][jk]}$. Denotaremos as coordenadas estandardizadas das colunas como Γ_{jka} . Agora, na AC, os resíduos de $[1][23]$ são decompostos, e as coordenadas são centradas de modo análogo a (56), i.e.:

$$\sum_{jk} p_{[ijk]} \Gamma_{jka} = 0 ,$$

i.e. para cada dimensão, as coordenadas têm média zero. Isto é devido ao fato de que as marginais são iguais (i.e. $p_{i[1..]} = e_{i[1..]} = 0$) (veja seção 2.5).

Não trataremos aqui o caso de um modelo de independência condicional ($[23][13]$, por exemplo).

Conclusões

- 10) Para tabelas multi-entrada, os resultados da AC aplicada a tabelas múltiplas, resultam na solução mostrando a diferença entre dois modelos log-lineares. Os resultados também mostram que AC não dá cobertura a todas as formas de dependência na matriz dos dados. Por exemplo, para a tabela múltipla

$P^{[l,jk]}$, a interação de primeira ordem entre as variáveis 2 e 3 não é mostrada.

- 20) O primeiro ponto motiva o seguinte: podemos usar estes resultados para combinar a AC e ALL. Por exemplo, quando a tabela múltipla mais adequada deve ser escolhida para analisar a tabela multi-entrada. Consideremos novamente a tabela de três entradas. Em primeiro lugar, se alguma interação de primeira ordem não é importante para nosso estudo, as duas variáveis correspondentes devem ser codificadas interativamente. Em segundo lugar, quando estamos especialmente interessados na relação de uma das variáveis com as outras duas, esta variável, não deve ser codificada interactivamente com outra variável. A AC simples mostra as interações de primeira ordem desta variável com as outras duas e a interação de segunda ordem. A generalização da AC, pode ser usada para suprimir outras interações de primeira ordem.
- 30) Outra implicação é que, em alguns casos, é possível usar AC e ALL de forma complementar. Van der Heijden e de Leeuw (1985) e van der Heijden, de Falguerolles e de Leeuw (1989) propuseram usar as duas técnicas para encontrar respostas a diferentes questões. Em um primeiro passo, ALL pode responder a questão de quais variáveis estão relacionadas, detectando as interações importantes; no segundo passo, a AC pode ser usada de duas maneiras: a) Quando um modelo no qual estamos particularmente interessados, não ajusta bem, AC pode ser usada para estudar os resíduos. b) Escolhendo um modelo que mal se ajusta, a AC pode ser usada para interpretar os parâmetros de interação. Pensa-se que em ambos os contextos, AC será especificamente útil quando a matriz sob estudo é grande, pois a representação gráfica pode simplificar a interpretação.

3.5 Análise de Correspondência para Modelos de Quase-Independência

Agora, usaremos a Análise de Correspondência Generalizada (ACG), para estudar o distanciamento da situação de independência em tabelas incompletas (veja van der Heijden (1986, 1987) e de Leeuw e van der Heijden (1988)). Os modelos log-lineares usados neste contexto são modelos de quase-independência. Neste trabalho, somente consideraremos tabelas incompletas de duas entradas, contudo, os resultados podem ser generalizados para modelos de independência para tabelas multi-entrada discutidas na seção 3.4 tais como fizemos para $[1||23]$, $[2||13]$ e $[3||12]$, ou da forma mais geral, modelos $[A||B]$ para grupos de categorias **A** e **B**. Os modelos de quase-independência são úteis em certas situações. Em primeiro lugar, quando algumas caselas da tabela são faltantes (missing), por exemplo elas estão disponíveis em algumas regiões (anos, grupos), mas não foram coletados em outras, apesar de que poderiam ter sido. Em segundo lugar, elas poderiam ser zeros estruturais (i.e. as caselas pertencem a fatos que logicamente não podem acontecer, por exemplo: as caselas diagonais de tabelas de importação-exportação). Em terceiro lugar, poderíamos optar por aplicar o modelo de independência a algumas caselas mas não a outras: as caselas restantes são observadas, mas não modeladas. Como um exemplo, em uma tabela de mobilidade social, poderia ser que estivessemos interessados somente nas caselas que refletem mudanças.

O modelo matemático é muito simples. Comparado com o modelo saturado (83), este modelo define independência somente para um número limitado de caselas, coletadas em **S**, enquanto que os valores para as caselas que estão fora de **S**, não têm restrições:

$$\begin{aligned} u_{12(ij)} &= 0 & , (i,j) \in S \\ \pi_{ij} &= p_{ij} & , \text{ caso contrário.} \end{aligned} \tag{88}$$

Este modelo pode ser testado com $(I-1)(J-1)-(IJ-\#S)$ graus de liberdade. Quando o distanciamento do modelo (88) é significativo, faz sentido estudar os resíduos deste modelo usando a ACG.

Começando da generalização da AC, como está definida na equação (82), tomaremos $\mathbf{Q}$ como a matriz de proporções do modelo de quase-independência. Para a escolha de $\mathbf{r}_g$ e $\mathbf{c}_g$, primeiro formulamos π_{ij} do modelo (88), como $\pi_{ij} = a_i b_j$, se a casela (i, j) está em $\mathbf{S}$. Independente da normalização escolhida (necessária para identificar o modelo), se tomamos a_i e b_j como os elementos de $\mathbf{r}_g$ e $\mathbf{c}_g$ respectivamente, então a soma dos valores singulares quadrados vem a ser:

$$\text{tr} (\mathbf{D}_\mu^2) = \sum_{\{(i,j) \in \mathbf{S}\}} \frac{(p_{ij} - \hat{\pi}_{ij})^2}{\hat{\pi}_{ij}} = \chi^2/n . \quad (89)$$

Note que, as caselas fora de $\mathbf{S}$ não são consideradas, pois não contribuem com a inércia.

Aqui, cada dimensão pode ser interpretada como a decomposição da estatística χ^2 , que é o critério usado para decidir se os resíduos contêm informação interessante.

Um esclarecimento sobre a escolha de $\mathbf{r}_g$ e $\mathbf{c}_g$, é o fato que nos leva a um procedimento de AC para tratamento de valores perdidos. Este procedimento foi proposto por Nora(1975) (veja também Greenacre(1984), pp. 236-244). Neste procedimento, os valores são atribuídos para as caselas perdidas a partir das caselas disponíveis da tabela. Estas atribuições têm a propriedade que, para uma dimensionalidade específica, digamos d , estes valores são perfeitamente representados pela solução da AC, é desta forma que o incremento da inércia devida às caselas perdidas, é nula.

Seguindo o modelo de quase-independência, se atribuímos valores "reconstituídos" para as caselas que não pertencem a $\mathbf{S}$, de tal forma que estas caselas não contribuam com resíduos, então a aplicação da AC decompõe os resíduos do modelo de quase-independência. A estratégia para achar tais atribuições está baseada no algoritmo E-M (Espectância-Maximização). Assim, escolhe-se primeiro a dimensionalidade d , logo, o processo iterativo começa com a solução ordinária da AC para a matriz, e usando a reconstrução de ordem d , atribuem-se iterativamente os valores para as caselas, i.e.: num passo h ($h = 1, 2, \dots$), atribue-se:

$$p_{(h)ij} = p_{ij}, \quad \text{se } (i,j) \in \mathbf{S} \quad (90)$$

$$p_{(h+1)ij} = p_{(h)i} \cdot p_{(h).j} (1 + \sum_a^d \mu_{(h)a} \phi_{(h)ia} \Gamma_{(h)ja}), \quad \text{caso contrário}$$

Este procedimento continua até que os valores estimados sejam praticamente os mesmos que no passo anterior.

Apesar que, o algoritmo converge em situações bem comportadas (i.e. não temos uma linha ou coluna totalmente missing), não temos segurança sobre a unicidade da estimativa resultante, ou da optimalidade dos gráficos. Mas, isto não é muito sério, pois o objetivo é poder usar a AC na tabela com todas suas linhas e colunas e não a atribuição optimal para as caselas perdidas. Este procedimento, elimina a indesejável alternativa de omitir linhas e colunas que contêm dados perdidos. Cabe esclarecer que, os eixos principais, não são aditivos como no caso de dados completos, pois a AC para $d = 2$, não necessariamente contém o eixo principal para $d = 1$, mas os eixos podem ser aproximadamente aditivos se os valores atribuídos são similares nas análises respectivas.

Em geral, a solução dependerá da dimensão d . Em particular, a reconstrução iterativa de ordem zero simplifica-se assim:

$$p_{(h)ij} = p_{ij}, \quad \text{se } (i,j) \in \mathbf{S} \quad (91)$$

$$p_{(h+1)ij} = p_{(h)i} \cdot p_{(h).j}, \quad \text{caso contrário.}$$

Isto é idêntico ao tratamento da AC para tabelas incompletas. Iterando, a equação (91) converge a valores independentes $a_i b_j$, calculadas para as caselas que não pertencem a $\mathbf{S}$. Logo, fixando o modelo de independência com os valores atribuídos $p_{(h)ij}$ para as caselas fora de $\mathbf{S}$, obteremos resíduos iguais a zero.

Daqui, deduzimos que, a AC para tabelas incompletas, pode ser calculada usando programas da AC clássicos. Somente temos que substituir os valores calculados com a equação (91). Isto faz com que este método seja fácil de interpretar: usando simplesmente AC aplicada à matriz $\mathbf{P}_{(h)}$.

Desta maneira, incluindo o modelo de quase-independência na abordagem da AC, temos ampliado o alcance da AC. E, um melhor entendimento da abordagem da AC para a reconstrução de ordem zero é obtida. O exemplo na

seção 4.3 mostra a aplicação da ACG para a decomposição dos resíduos do modelo de quase-independência.

3.6 Análise de Correspondência para Decompor Resíduos do Modelo de Simetria

Recentemente, muita atenção tem se dado à anti-simetria em tabelas quadradas. Uma importante linha de contribuições focaliza atenção na anti-simetria de uma matriz quadrada.

O modelo log-linear de simetria para uma matriz quadrada, se expressa assim:

$$\log \pi_{ij} = u + u_{1(i)} + u_{2(j)} + u_{12(ij)}, \quad (92)$$

onde:

$$u_{1(i)} = u_{2(i)}, \text{ e}$$

$$u_{12(ij)} = u_{12(ji)}.$$

Este modelo pode ser testado com $I(I-1)/2$ graus de liberdade.

As estimativas de máxima verossimilhança para este modelo, estão dados na matriz simétrica $\mathbf{S}_{IXI}$, onde:

$$s_{ij} = s_{ji} = (p_{ij} + p_{ji})/2. \quad (93)$$

Logo, matriz de resíduos do modelo de simetria $\mathbf{A}$, que chamaremos de *matriz de anti-simetria*, pode ser formulada pela expressão da matriz $\mathbf{P} - \mathbf{S}$, construída para refletir o desvio da simetria, assim:

$$\mathbf{P} = \mathbf{S} + \mathbf{A} \quad (94)$$

onde: $a_{ij} = -a_{ji} = (p_{ij} - p_{ji})/2$.

Gower (1977) e Constantine e Gower (1978), trabalharam sobre a anti-simetria e mostraram que a decomposição em valores singulares de uma matriz anti-simétrica $\mathbf{A}$ tem a forma:

$$\mathbf{A} = \mathbf{U} \mathbf{D}_\mu \mathbf{V}' = \mathbf{U} \mathbf{D}_\mu \mathbf{J} \mathbf{U}' , \quad (95)$$

onde: - $\mathbf{J}$ é uma matriz bloco diagonal com blocos:

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

- os valores singulares estão ordenados em pares

$$\mu_1, \mu_1, \mu_2, \mu_2, \dots, e$$

- se o número de linhas é ímpar, $\mu_I = 0$ e $\mathbf{J}_{II} = 0$.

Assim, quando fazemos um gráfico dos vetores singulares $\mathbf{U}$ e $\mathbf{UJ}$ para dimensões pareadas, a configuração dos pontos linha é a mesma dos pontos coluna, só que com uma rotação de 90° . Um gráfico mais simples pode ser feito colocando somente os pontos linha e dando a direção da rotação.

Gower (1977) e Constantine e Gower (1978), mostraram que, no gráfico dos pontos linha acontece o seguinte:

- Os elementos a_{ij} e a_{ji} , estão aproximados por duas vezes a área do triângulo formado pela linha i , linha j e a origem (este triângulo será denotado por O_{ij}).
- Quando a rotação dos pontos linha aos pontos coluna vai no sentido horário, o elemento a_{ij} é positivo se a rotação da linha i à linha j é no sentido anti-horário e a_{ji} é negativo se a rotação for no sentido horário.
- Quando a rotação dos pontos linha aos pontos coluna vai no sentido anti-horário, o elemento a_{ij} é positivo se a rotação da linha i à linha j é no sentido horário e a_{ji} é negativo se a rotação for no sentido anti-horário.
- Os pontos sobre uma linha que passa pela origem não têm relação de anti-simetria porque a área O_{ij} é zero. Os pontos que estão perto da origem têm anti-simetria pequena.

Cabe esclarecer que, estas relações são satisfeitas quando duas dimensões dão uma boa aproximação de $\mathbf{A}$.

Vejamos a relação que tem a AC com a decomposição da *anti-simetria* dada em (95): esta decomposição é um caso particular da ACG (compare com a equação (82)), na qual tomamos $\mathbf{O} = \mathbf{S}$, e $\mathbf{r}_r = \mathbf{c}_r = \mathbf{1}_I$ (com o que não eliminamos a influência das marginais). Na abordagem da AC, é regra ponderar linhas e colunas, para eliminar a influência das marginais da tabela; a ponderação da tabela dos resíduos é também necessária neste contexto. Com o propósito de manter as propriedades da decomposição em valores singulares da desviação da simetria, as ponderações das linhas devem ser iguais as das colunas: $\mathbf{r}_r = \mathbf{c}_r = \mathbf{r}_0 = \mathbf{c}_0 = (\mathbf{r} + \mathbf{c})/2$. Logo, com a DVS da matriz $\mathbf{D}_{rr}^{-1/2} \mathbf{A} \mathbf{D}_{rr}^{-1/2}$, a decomposição da AC fica assim:

$$\mathbf{P} = \mathbf{S} + \mathbf{D}_{rr} \Phi \mathbf{D}_p \mathbf{J} \Phi' \mathbf{D}_{rr}, \quad (96)$$

Como as marginais de $\mathbf{P}$ não são iguais às de $\mathbf{O}$, é possível encontrar a situação em que todos os pontos estejam a um lado do espaço (veja seção 2.5). O exemplo 4.3 mostra a aplicação da ACG para a decomposição dos resíduos do modelo de simetria.

3.7 Análise de Correspondência para Decompor Resíduos do Modelo de Quase-Simetria

Nos resíduos do modelo de simetria, a anti-simetria é em parte devida às diferenças entre as marginais de $\mathbf{P}$ e em parte devida a outros aspectos. Raramente, conseguimos bom ajuste com o modelo de simetria, pois ele é altamente estruturado. O modelo de *quase-simetria* (Caussinus, 1965; Bishop et al., 1975) é um modelo com o qual é possível estudar se a anti-simetria, que não se deve às marginais diferentes, pode ser desprezada. Este modelo pode ser formulado como:

$$\log \pi_{i,j} = u + u_{1(i)} + u_{2(j)} + u_{12(i,j)}, \quad (97)$$

com a restrição:

$$u_{12(ij)} = u_{12(ji)}.$$

Este modelo tem $(I-1)(I-2)/2$ graus de liberdade. Comparado com o modelo de simetria (92), a restrição $u_{1(i)} = u_{2(i)}$ não é considerada. Se $p_{i.} = p_{.i}$, o modelo de quase-simetria fica simplificado como o modelo de simetria.

O modelo de quase-simetria, não requer homogeneidade marginal, como o modelo de simetria. As estimativas de máxima verossimilhança para este modelo, estão dadas na matriz quase-simétrica $\mathbf{T}_{IXI}$, onde:

$$t_{i.} = p_{i.},$$

$$t_{.j} = p_{.j}.$$

$$t_{ij} + t_{ji} = p_{ij} + p_{ji}, \quad (i, j = 1, \dots, I).$$

somente necessitamos de uma das duas primeiras equações; dada a terceira, uma das primeiras é redundante.

A matriz $\mathbf{T}$, não é necessariamente simétrica; entanto que a matriz residual $\mathbf{A} = \mathbf{P} - \mathbf{T}$ é uma matriz anti-simétrica (pois, $a_{ij} = -a_{ji}$). Queremos estudar a matriz residual $\mathbf{A}$ com AC, mantendo as propriedades da DVS da anti-simetria. De forma análoga à análise para os resíduos do modelo de simetria, isto pode ser efetuado tomando o mesmo conjunto de pesos para as linhas e as colunas, de modo que novamente tomamos $(p_{i.} + p_{.i})/2$ como os elementos de $\mathbf{r}_g$ e $\mathbf{c}_g$. Agora temos a mesma situação que na seção anterior. A fórmula de reconstrução pode ser escrita como na equação (96), com a diferença que $\mathbf{Q} = \mathbf{T}$ não é simétrica.

3.8 Observações Gerais

Nesta seção, temos apresentado algumas relações úteis entre AC e ALL:

1. Em primeiro lugar, os scores obtidos com AC, usualmente se aproximam aos parâmetros do modelo de associação RC de Goodman.
2. Em segundo lugar, temos visto que AC de tabelas multi-dimensionais por meio de tabelas múltiplas, pode ser interpretada em termos de alguns modelos log-lineares de independência, e temos indicado que AC pode guiar a interpretação dos parâmetros de interação ou dos resíduos.
3. Em terceiro lugar, o procedimento da AC para o tratamento de caselas perdidas está estreitamente relacionado ao modelo de quase-independência.
4. Temos demonstrado como AC e ACG podem ser usadas para a análise de tabelas quadradas. A AC pode nos fornecer uma clara visão das anti-simetrias importantes de uma matriz quadrada.

4. Alguns Exemplos de Aplicação

4.1 Aplicação da AC a um Exemplo Simples de Lingüística (Matilde Scaramucci, IEL/UNICAMP, 1993)

Os dados da tabela 4.1.1 foram obtidos da Pesquisa: "*Papel do Léxico na Leitura em Inglês como Língua Estrangeira: Foco no Produto e no Processo*", feita por Matilde Scaramucci, aluna do Doutorado em Lingüística IEL/UNICAMP (Instituto de Estudos da Linguagem) e professora de inglês do DLA/UNICAMP (Departamento de Lingüística Aplicada). Os indivíduos, dos quais os dados foram coletados, foram alunos do curso de Inglês Instrumental, oferecido pelo CEL/UNICAMP, durante o segundo semestre de 1990. Estes alunos, são predominantemente alunos regulares da UNICAMP, das áreas de Matemática Aplicada e Engenharia Elétrica.

Entre outros, foi avaliado um teste de familiaridade para um grupo de 30 palavras, sendo 20 reais e 10 inventadas (veja a tabela 4.1.1). Neste teste, não foi advertido a existência de palavras inventadas; cabe esclarecer que, entre as palavras reais, existem palavras "cognatas" (mesma origem em duas línguas, mesmos radicais) e também, existem palavras "falsas cognatas" (parecem, mas não são o que parecem), por outro lado, para as palavras inventadas, existem algumas "falsas cognatas". As classificações para o teste de familiaridade, foram:

- a palavra é *não familiar* (não conheço o significado).
- a palavra é *pouco familiar* (não tenho certeza sobre o significado).
- a palavra é *familiar* (sei o significado)

Devido a que a tabela 4.1.1 tem 30 linhas x 3 colunas, os dados serão completamente representados em gráficos de duas dimensões (i.e.: a qualidade de representação será de 100% para todos os pontos em duas dimensões). Da tabela 4.1.3, temos que o primeiro eixo explica 83% da inércia, e o segundo 17%.

Da tabela 4.1.4, as categorias coluna que têm maior contribuição absoluta sobre o primeiro eixo são: *familiar* do lado positivo com uma contribuição igual a 61% e *não familiar* do lado negativo com uma contribuição absoluta igual a 38%, de modo que, estas duas categorias de familiaridade explicam o 99% da inércia do primeiro eixo. Ao mesmo tempo, estes pontos têm uma excelente qualidade de representação pelo primeiro eixo e são 96% e 97% respectivamente.

Portanto, este eixo vai, no sentido crescente, de menor a maior grau de familiaridade e pode ser interpretado como um fator de familiaridade. Olhando para o segundo eixo, a maior contribuição absoluta está dada pela categoria *pouco familiar* do lado positivo, com uma contribuição absoluta igual a 83% e uma qualidade de representação 92%, conseqüentemente, o lado positivo do eixo pode ser interpretado como familiaridade média; se alguma palavra estivesse bem representado do lado negativo, a palavra teve altas classificações como *familiar* e *não familiar*, mas baixa classificação como *pouco familiar*. Com estes dois resultados, vemos que o grau de familiaridade segue um "arco" no espaço, este fenômeno é chamado de "*Efeito de Guttman*" e, se dá quando existe uma ordem entre as categorias segundo os dados da tabela. Conseqüentemente, das apreciações a respeito dos eixos, as posições das palavras no espaço indicam o grau de familiaridade segundo o grupo de alunos que participaram do teste.

Com o propósito de ilustrar a comparação entre o gráfico em coordenadas principais e o gráfico combinando ambos os tipos de coordenadas, no gráfico 4.1.1, foram colocadas as palavras e categorias de familiaridade em coordenadas principais e, em adição, colocamos no mesmo gráfico as posições das categorias de familiaridade em coordenadas standardizadas. Ao inspecionar tal gráfico, podemos dizer rapidamente:

- As palavras "mais familiares" foram (em ordem de maior a menor grau de familiaridade): close, slowly, empty, spring, receiving, e appeared.
- As palavras de "familiaridade média" foram (em ordem de maior a menor grau de familiaridade): cloud, *adjustion, picked up, pattern, shaken, *observement, throughout, massive, fairly, fortuitously e disk-shaped.
- Finalmente, as palavras menos familiares foram (em ordem de menor a maior grau de familiaridade): *drealled, *glim, *tark, *trivel, loomed, spawned, *lodings, *leam, resembling, *resks, *skapping, cluster e focusing.

Observe que este gráfico revela a informação da tabela com bom detalhe, por exemplo: As palavras inventadas foram classificadas, pelos alunos, como não familiares, que é justamente o que esperávamos, a excessão das palavras *adjustion e *observement que apesar de serem inventadas, foram classificadas

como pouco familiares talvez, por serem falsas cognatas (pois foram construídas com radicais e finais existentes, embora a combinação não exista como palavra) e podiam confundir aos alunos. Por outro lado, a palavra throughout foi mais classificada como pouco familiar que qualquer outra palavra. No caso das palavras massive e fortuitously, estão situadas bem perto da origem, de modo que têm um perfil similar ao perfil médio (confira na tabela de perfis linha 4.1.2).

Vemos então que, a AC grafica a tabela, de forma que a informação contida na tabela é fácil e rápida de entender, este gráfico pode ser interpretado como aquele onde estão contidas simultaneamente as informações de todos os histogramas de cada palavra.

Tabela 4.1.1: Tabela de Contingência para a Classificação de Palavras segundo o teste de Familiaridade (as palavras inventadas estão precedidas por um *).

Palavra	Familiaridade			Total
	<i>não familiar</i>	<i>pouco familiar</i>	<i>familiar</i>	
*lodings	45	3	1	49
cluster	37	7	5	49
*skapping	41	6	2	49
spawned	45	4	0	49
shaken	21	7	21	49
*glim	47	2	0	49
loomed	47	1	1	49
*drealled	48	1	0	49
resembling	40	8	1	49
cloud	18	10	21	49
*leam	44	4	1	49
disk-shaped	27	15	7	49
*tark	47	0	1	48
fortuitously	27	10	12	49
fairly	21	17	11	49
*observement	16	15	16	47
massive	26	7	16	49
*trivel	45	3	0	48
picked up	17	13	19	49
focusing	34	9	6	49
spring	9	5	35	49
empty	6	4	39	49
throughout	15	21	13	49
pattern	17	14	18	49
appeared	14	10	25	49
close	3	2	44	49
*adjustion	15	16	18	49
receiving	6	12	31	49
*resks	40	8	1	49
slowly	6	1	42	49
Total	824	235	407	1466

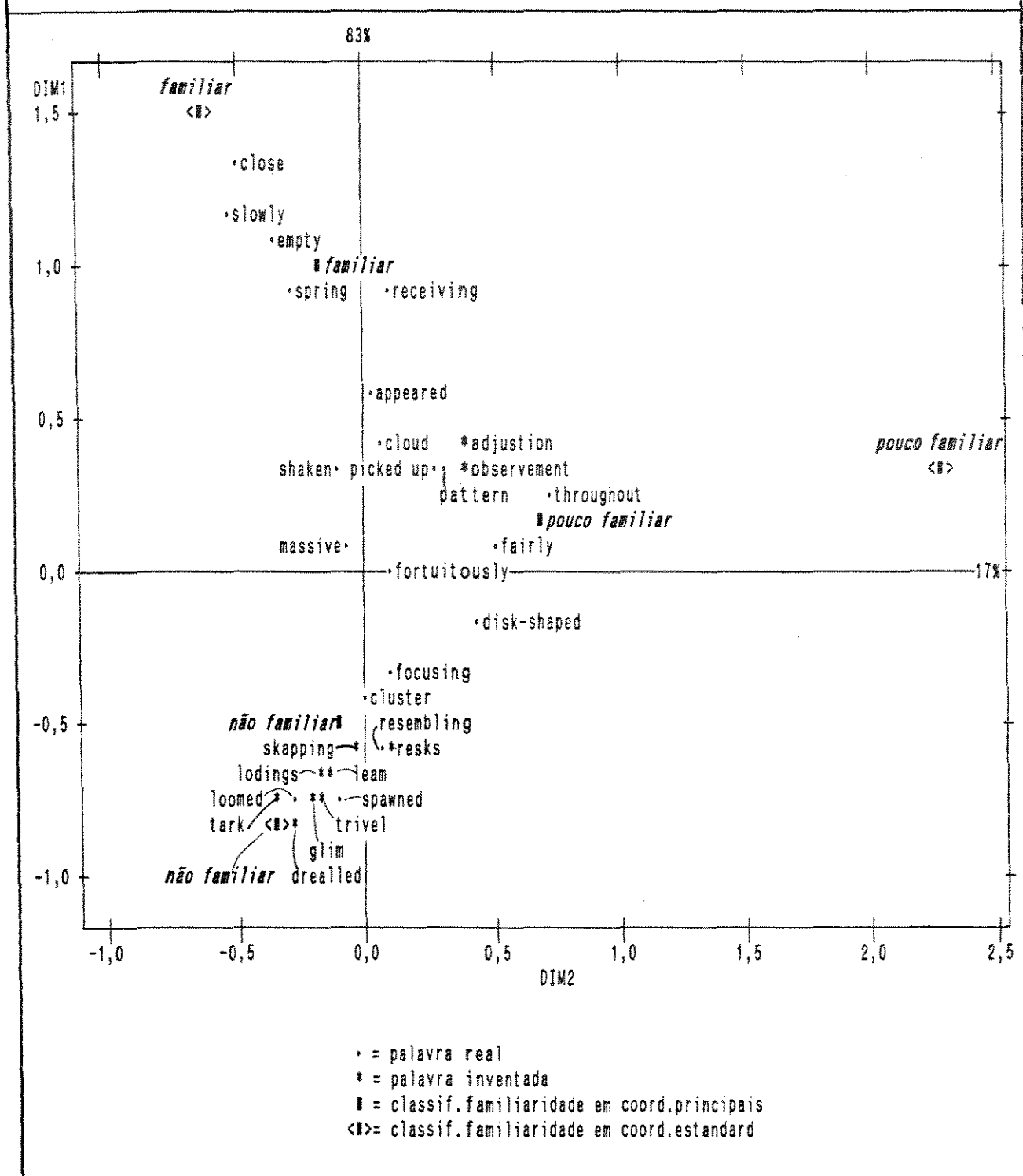
Fonte: Matilde Scaramucci, IEL/UNICAMP, 1993.

Tabela 4.1.2 : Perfis Linha (* 100)			
Palavra	Familiaridade		
	<i>não familiar</i>	<i>pouco familiar</i>	<i>familiar</i>
*lodings	91,84	6,12	2,04
cluster	75,51	14,29	10,20
*skapping	83,67	12,24	4,08
spawned	91,84	8,16	0,00
shaken	42,86	14,29	42,86
*glim	95,92	4,08	0,00
loomed	95,92	2,04	2,04
*drealled	97,96	2,04	0,00
resembling	81,63	16,33	2,04
cloud	36,73	20,41	42,86
*learn	89,80	8,16	2,04
disk-shaped	55,10	30,61	14,29
*tark	97,92	0,00	2,08
fortuitously	55,10	20,41	24,49
fairly	42,86	34,69	22,45
*observement	34,04	31,91	34,04
massive	53,06	14,29	32,65
*trivel	93,75	6,25	0,00
picked up	34,69	26,53	38,78
focusing	69,39	18,37	12,24
spring	18,37	10,20	71,43
empty	12,24	8,16	79,59
throughout	30,61	42,86	26,53
pattern	34,69	28,57	36,73
appeared	28,57	20,41	51,02
close	6,12	4,08	89,80
*adjustion	30,61	32,65	36,73
receiving	12,24	24,49	63,27
*resks	81,63	16,33	2,04
slowly	12,24	2,04	85,71
Perfil Médio	56,21	16,03	27,76

Tabela 4.1.3 : Decomposição da Inércia relativo aos Eixos									
Eixo	Valores Singul.	Inércias Princip.	Chi-Quadr.	Percent	17	34	51	68	85
					-----+-----+-----+-----+-----				
1º	0,65757	0,43240	633,898	83,15%	*****				
2º	0,29605	0,08765	128,491	16,85%	****				
		-----	-----						
Total		0,52005	762,389 (Graus de Liberdade = 58)						

Tabela 4.1.4 : Estatísticas para as Classificações de Familiaridade e as Palavras								
Classif.Famil. /Palavras	Massa	Proporção Inércia	Coord.Principais		Contrib.Absoluta		Contrib.Relativa	
			1º Eixo	2º Eixo	1º Eixo	2º Eixo	1º Eixo	2º Eixo
<i>não familiar</i>	0,56207	0,32246	-0,53717	-0,09899	0,37508	0,06284	0,96715	0,03285
		(CE)(-0,81690	-0,33437)					
<i>pouco familiar</i>	0,16030	0,15078	0,19408	0,67193	0,01396	0,82574	0,07700	0,92300
		(CE)(0,29514	2,26962)					
<i>familiar</i>	0,27763	0,52676	0,97548	-0,18755	0,61095	0,11142	0,96435	0,03565
		(CE)(1,48345	-0,63351)					
*loadings	0.03342	0.03377	-0.70187	-0.18105	0.03808	0.01250	0.93761	0.06239
cluster	0.03342	0.01152	-0.42331	0.00710	0.01385	0.00002	0.99972	0.00028
*skipping	0.03342	0.02218	-0.58684	-0.02773	0.02662	0.00029	0.99777	0.00223
spawned	0.03342	0.03484	-0.72612	-0.12180	0.04076	0.00566	0.97263	0.02737
shaken	0.03342	0.00743	0.32783	-0.09057	0.00831	0.00313	0.92908	0.07092
*glim	0.03342	0.04160	-0.77151	-0.22809	0.04601	0.01984	0.91962	0.08038
loomed	0.03342	0.04120	-0.74726	-0.28733	0.04316	0.03148	0.87119	0.12881
*drealled	0.03342	0.04562	-0.79420	-0.28123	0.04876	0.03016	0.88858	0.11142
resembling	0.03342	0.02271	-0.58839	0.08466	0.02676	0.00273	0.97972	0.02028
cloud	0.03342	0.01038	0.39591	0.06885	0.01212	0.00181	0.97064	0.02936
*leam	0.03342	0.03070	-0.67917	-0.12791	0.03566	0.00624	0.96575	0.03425
disk-shaped	0.03342	0.01274	-0.14786	0.42004	0.00169	0.06728	0.11025	0.88975
*tark	0.03274	0.04453	-0.76897	-0.34060	0.04478	0.04334	0.83599	0.16401
fortuitously	0.03342	0.00103	-0.02660	0.12380	0.00005	0.00584	0.04413	0.95587
fairly	0.03342	0.01666	0.08532	0.50190	0.00056	0.09606	0.02808	0.97192
*observement	0.03206	0.01597	0.32111	0.39486	0.00764	0.05703	0.39807	0.60193
massive	0.03342	0.00079	0.09310	-0.06005	0.00067	0.00138	0.70619	0.29381
*trivel	0.03274	0.03702	-0.74739	-0.17162	0.04230	0.01100	0.94991	0.05009
picked up	0.03342	0.01252	0.37011	0.24049	0.01059	0.02206	0.70312	0.29688
focusing	0.03342	0.00778	-0.33097	0.10728	0.00847	0.00439	0.90492	0.09508
spring	0.03342	0.06188	0.93968	-0.28233	0.06826	0.03040	0.91720	0.08280
empty	0.03342	0.08677	1.10477	-0.35989	0.09435	0.04939	0.90406	0.09594
throughout	0.03342	0.03638	0.26999	0.70226	0.00563	0.18807	0.12877	0.87123
pattern	0.03342	0.01346	0.34585	0.29974	0.00925	0.03426	0.57107	0.42893
appeared	0.03342	0.02202	0.58370	0.04443	0.02634	0.00075	0.99424	0.00576
close	0.03342	0.12349	1.29411	-0.49670	0.12946	0.09408	0.87160	0.12840
*adjustion	0.03342	0.02043	0.39124	0.40602	0.01183	0.06287	0.48147	0.51853
receiving	0.03342	0.05415	0.91076	0.11409	0.06412	0.00496	0.98455	0.01545
*resks	0.03342	0.02271	-0.58839	0.08466	0.02676	0.00273	0.97972	0.02028
slowly	0.03342	0.10769	1.17753	-0.53763	0.10718	0.11023	0.82750	0.17250

Gráfico 4.1.1 : As palavras e as classificações de Familiaridade em coordenadas principais.



4.2 Aplicação da AC como um suplemento da aplicação da ALL a um Exemplo de Pena de Morte (Radelet, 1981)

Este exemplo, foi extraído do livro de Agresti (1990). Dos dados, temos as seguintes informações:

A tabela 4.2.1, foi apresentada por Radelet (1981), é uma tabela tri-dimensional 2x2x2. O artigo de Radelet, estudou os efeitos de características raciais sobre a pena de morte para acusados de homicídio. As variáveis da tabela 4.2.1, são:

- Veredicto de Pena de Morte, com categorias: Condenação e Absolvição.
- Raça do Réu, com as categorias: Branca e Negra.
- Raça da Vítima, com as categorias: Branca e Negra

Os sujeitos foram réus de homicídio em 20 municípios do Estado da Flórida dos Estados Unidos durante os anos 1976 a 1977. O número total de réus foi de 326.

Tabela 4.2.1: Veredicto da Pena de Morte segundo a Raça do Réu e a Raça da Vítima.			
Raça do Réu	Raça da Vítima	Pena de Morte	
		Sim	Não
Branca	Branca	19	132
	Negra	0	9
Negra	Branca	11	52
	Negra	6	97

Fonte: Agresti(1990), p 136.; Radelet(1981).

Primeiro faremos as análises das três tabelas múltiplas, construídas codificando duas variáveis interactivamente contra a terceira e depois faremos

as análises correspondentes às três tabelas marginais. Ambas as análises serão feitas, pois veremos que o modelo que se ajusta é o modelo de independência condicional, tal modelo não implica o modelo de independência marginal (veja Agresti, 1990, p.138-142.)

4.2.1 Análise da Tabela

Segundo a tabela 4.2.2 temos que o modelo que se ajusta bem aos dados da tabela de três entradas é o modelo [12][23], isto é:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{23(jk)},$$

Tabela 4.2.2: Tabela da Análise de Variância para o Modelo [12][23]			
Fonte	GL	Chi-Quadrado	Prob
RÉU	1	15.76	0.0001
VITIMA	1	29.50	0.0000
VEREDICTO	1	97.38	0.0000
RÉU*VITIMA	1	73.28	0.0000
VITIMA*VEREDICTO	1	4.45	0.0348
RESIDUAL	1	0.87	0.3509

conseqüentemente, as interações desprezíveis são: a de primeira ordem Réu*Veredicto e a de segunda ordem Réu*Vítima*Veredicto. Para interpretarmos a interação Réu*Vítima, deveremos aplicar AC a tabela múltipla construída codificando interactivamente Vítima com Veredicto e, para interpretarmos a interação Vítima*Veredicto, deveremos aplicar AC a tabela múltipla construída codificando interactivamente Réu com Vítima.

Interpretando a Interação Réu*Vítima

A análise de correspondência da tabela múltipla construída codificando interactivamente Vítima com Veredicto, decompõe os resíduos do modelo de independência conjunta [1][23], i.e. o modelo:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)},$$

visto que, aquele que bem se ajusta é o modelo [12][23], então esta análise decompõe a interação Réu*Vítima.

Esta tabela pode ser representada 100% num espaço de uma dimensão. O gráfico 4.2.1 revela as tendências dos réus dada alguma combinação vítima*veredicto:

- a. O réu foi maiormente de raça branca, em primeiro lugar, quando a vítima era branca e o veredicto foi de absolvição e, em segundo lugar, quando a vítima era branca e o veredicto foi a condenação.
- b. O réu foi maiormente de raça negra, em primeiro lugar, quando a vítima era negra e houve condenação e, em segundo lugar, quando a vítima era negra e não houve condenação. Mais ainda, o gráfico 4.2.1.b revela que em todos os casos em que a vítima era negra e houve pena de morte, o réu era negro.

As observações em (a) e (b) indicam que as vítimas foram mortas principalmente por réus da mesma raça, assim podemos interpretar que a interação Réu*Vítima se dá por que os homicídios foram feitos por pessoas da mesma raça que a vítima.

Interpretando a Interação Vítima*Veredicto

A análise de correspondência da tabela múltipla construída codificando interactivamente Réu com Vítima, decompõe os resíduos do modelo de independência conjunta [3][12], i.e. o modelo:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)},$$

assim, esta análise decompõe a interação Vítima*Veredicto.

Esta tabela pode ser representada 100% num espaço de uma dimensão. O gráfico 4.2.2 revela as prioridades que foram dadas nos veredictos dos julgamentos na Flórida:

- a. As maiores chances do réu ser condenado a pena de morte foram, em primeiro lugar, quando o réu era negro e a vítima branca e, em segundo lugar, quando o réu era branco e a vítima branca.

- b. As maiores chances do réu não ser condenado a pena de morte foram, em primeiro lugar, quando o réu era branco e a vítima negra e, em segundo lugar, quando o réu era negro e a vítima negra. Mais ainda, o gráfico 4.2.2.b revela que não houve nenhuma condenação quando o réu foi branco e a vítima negra.

As observações em (a) e (b) indicam que o veredito dependeu principalmente da vítima, quando a vítima era branca houveram mais condenações a morte que quando a vítima era negra.

Interpretando a Soma das Interações Réu*Vítima e Vítima*Veredicto

Nas análises anteriores, temos indicado interpretações das interações Réu*Vítima e Vítima*Veredicto. Podemos ganhar ainda mais quando aplicarmos a análise de correspondência à tabela múltipla construída codificando interactivamente Réu com Veredicto, para decompor os resíduos do modelo de independência conjunta [2][13], i.e. o modelo:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{13(ik)},$$

visto que o modelo que bem se ajusta é o modelo [12][23], então esta análise decompõe a soma das interações Réu*Vítima e Vítima*Veredicto.

No gráfico 4.2.3, vemos que a direção da interação Réu*Vítima se manteve igual a da análise anterior, entanto que a direção da interação Vítima*Veredicto não se manteve, isto revela que a interação Réu*Vítima foi mais forte que a de Vítima*Veredicto (claro, isto poderia ser concluído olhando diretamente a tabela 4.2.2, ou às estatísticas Chi-Quadrado da tabela 4.2.3). O gráfico 4.2.3.b, revela, o que já conhecíamos, só que expressado de outro modo, que em todos os casos em que o réu foi branco e ele condenado, a vítima foi branca.

Tabela 4.2.3 : Estatísticas para cada Tabela Múltipla Combinando duas Variáveis Contra a Terceira									
Tabela Múltipla	Valores Inércia Chi-Singul. Princip. Quadr.			Categoria	Massa	Proporção Inércia	Coord. Princip.	Coord. Estand.	Contrib Absolut
Vit*Vered x Réu	0.59626	0.35553	115.901 (3g1)	V bran * Condena	0.09202	0.02104	0.28512		0.02104
				V bran * Não con	0.56442	0.32616	0.45326		0.32616
				V negr * Condena	0.01840	0.04990	-0.98176		0.04990
				V negr * Não con	0.32515	0.60290	-0.81192		0.60290
				R bran	0.49080	0.50920	0.60734	1.01858	0.50920
				R negr	0.50920	0.49080	-0.58538	-0.98176	0.49080
Réu*Vitim x Veredicto	0.14630	0.02140	6.97734 (3g1)	R bran * V bran	0.46319	0.05224	0.04913		0.05224
				R bran * V negr	0.02761	0.16012	-0.35233		0.16012
				R negr * V bran	0.19325	0.37853	0.20475		0.37853
				R negr * V negr	0.31595	0.40911	-0.16647		0.40911
				Condenado	0.11043	0.88957	0.41523	2.83823	0.88957
				Não conde	0.88957	0.11043	-0.05155	-0.35233	0.11043
Réu*Vered x Vítima	0.61004	0.37215	121.322 (3g1)	R bran * Condena	0.05828	0.08196	-0.72344		0.08196
				R bran * Não con	0.43252	0.40323	-0.58903		0.40323
				R negr * Condena	0.05215	0.00005	0.01976		0.00005
				R negr * Não con	0.45706	0.51475	0.64740		0.51475
				V bran	0.65644	0.34356	-0.44133	-0.72344	0.34356
				V negr	0.34356	0.65644	0.84325	1.38229	0.65644

Gráfico 4.2.1 : Tabela Combinando Vítima e Veredito contra Réu (revela interação Réu*Vítima)

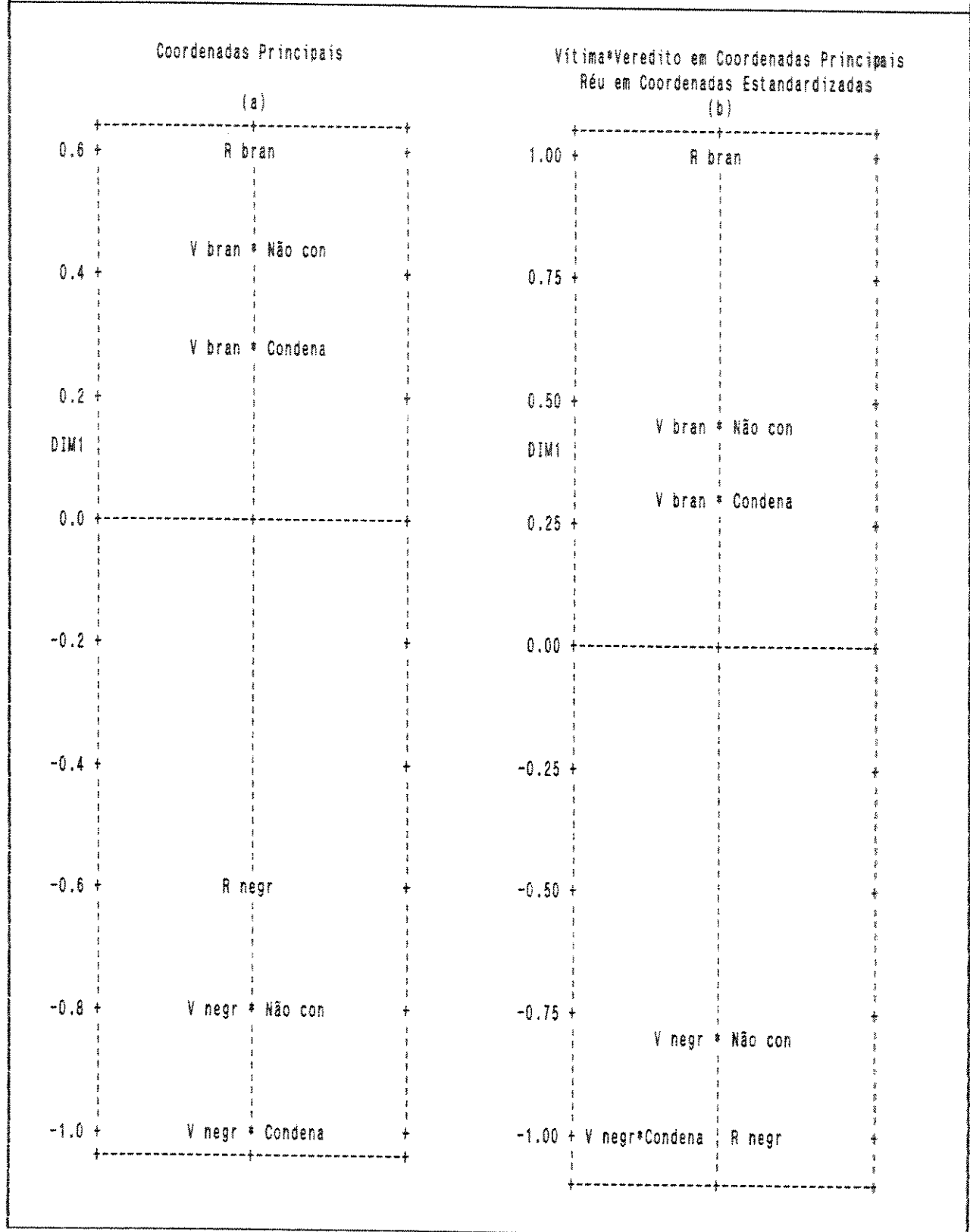


Gráfico 4.2.2 : Tabela Combinando Réu e Vítima contra Veredito (revela interação Vítima*Veredito)

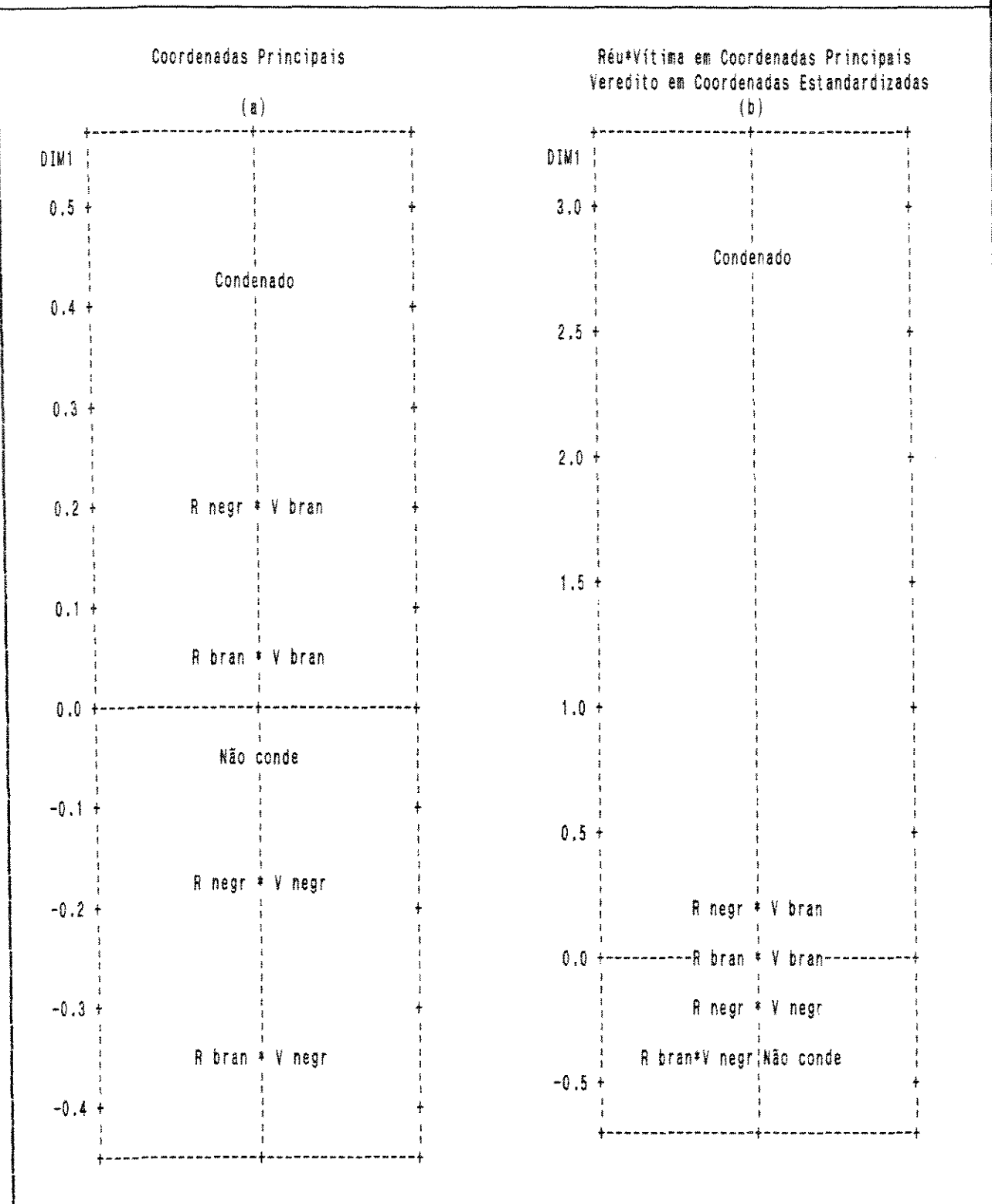
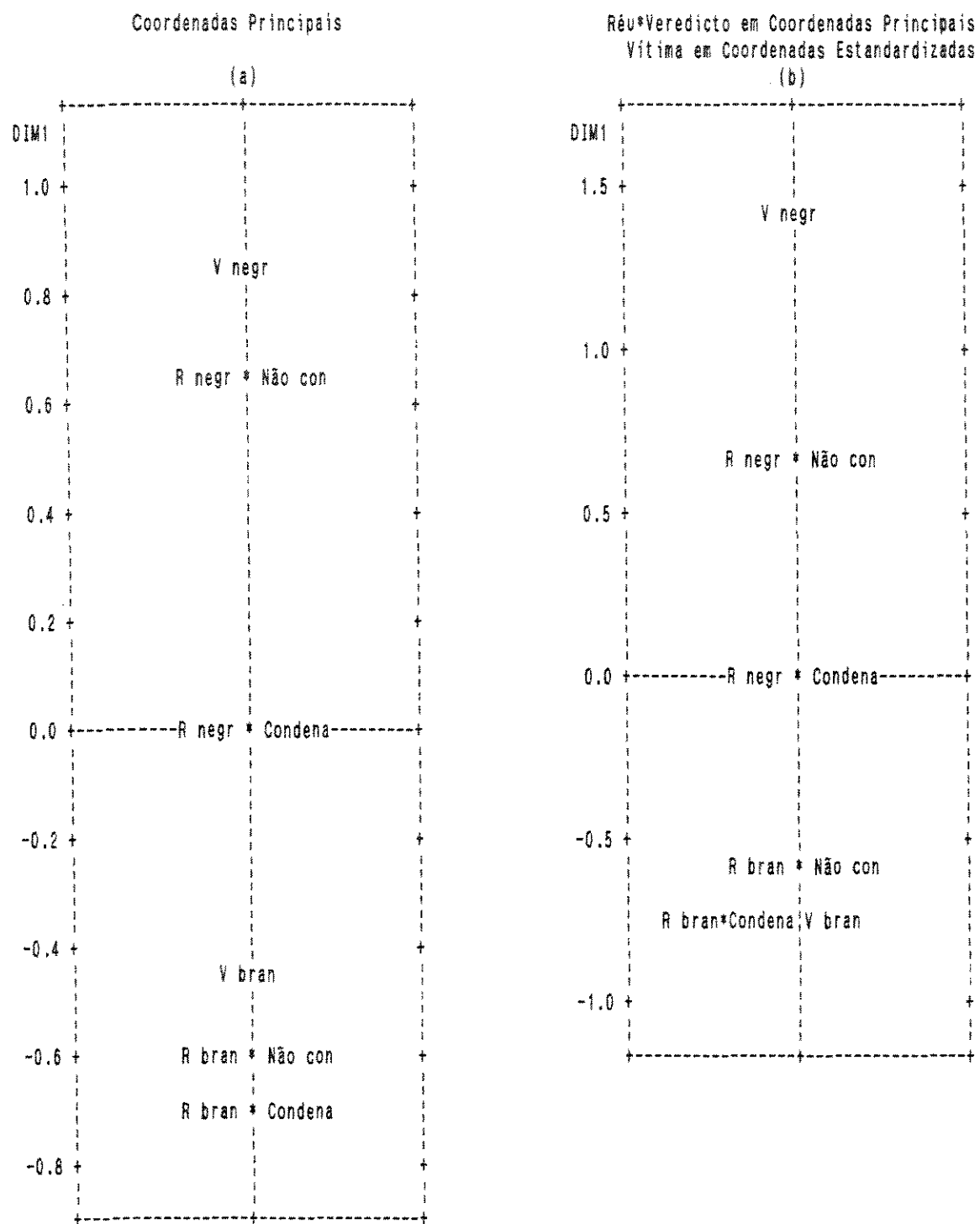


Gráfico 4.2.3 : Tabela Combinando Réu e Veredicto contra Vítima (revela a soma das interações Réu*Vítima e Vítima*Veredicto)



4.2.2 Análises das Tabelas Marginais

Interpretando a Interação Réu*Vítima da Tabela Marginal

Segundo a tabela 4.2.4, temos que o modelo de independência marginal [1][2] se ajusta muito mal aos dados.

Tabela 4.2.4: Tabela da Análise de Variância para o Modelo Marginal [12]			
Fonte	GL	Chi-Quadrado	Prob
RÉU	1	17.05	0.0000
VITIMA	1	37.83	0.0000
RÉU*VITIMA	1	76.52	0.0000

A análise de correspondência da tabela marginal Réu Vítima, decompõe a interação Réu*Vítima da tabela marginal. O gráfico 4.2.4 revela que a tendência é que o réu ataca principalmente vítimas da própria raça.

Interpretando os Resíduos do modelo de Independência da Tabela Marginal Réu e Veredicto

Segundo a tabela 4.2.5, temos que o modelo de independência marginal [1][3] se ajusta bem aos dados.

Tabela 4.2.5: Tabela da Análise de Variância para o Modelo Marginal [1][3]			
Fonte	DF	Chi-Square	Prob
RÉU	1	0.11	0.7395
VEREDICTO	1	138.95	0.0000
RESIDUAL	1	0.22	0.6382

A análise de correspondência da tabela marginal Réu Veredicto, decompõe os resíduos do modelo de independência. Mesmo que a interação Réu*Veredicto tenha sido não significativa, podemos conhecer a direção desta, ao examinarmos o gráfico 4.2.5.c que revela que dentro do conjunto de condenados, os réus eram na maioria brancos.

É interessante comparar este resultado com a análise feita para a tabela múltipla combinando Réu e Vítima, que indica que a condenação depende maiormente da vítima ser branca e o réu negro. Isto, mostra a importância das análises anteriores da tabela múlti-entradas.

Interpretando a Interação Vítima*Veredicto da Tabela Marginal

Segundo a tabela 4.2.6, temos que o modelo de independência marginal [2][3] se ajusta mal aos dados.

Tabela 4.2.6): Tabela da Análise de Variância para o Modelo Marginal [23]			
Fonte	GL	Chi-Quadrad	Prob
VITIMA	1	21.73	0.0000
VEREDICTO	1	102.17	0.0000
VITIMA*VEREDICTO	1	5.21	0.0225

A análise de correspondência da tabela marginal Vítima Veredicto, decompõe a interação Vítima*Veredicto da tabela marginal. O gráfico 4.2.4.c dentro do conjunto dos julgamentos com condenação, a maior proporção está dada quando a vítima é branca.

Tabela 4.2.7 : Estatísticas para cada Tabela Marginal da tabela 4.2. 1									
Tabela Marginal	Valores Singul.	Inércia Princip.	Chi-Quadr.	Categoria	Massa	Proporção Inércia	Coord. Princip.	Coord. Estand.	Contrib Absolut
Réu x Vítima	0.59396	0.35279	115.008 (1g1)	R bran	0.49080	0.50920	-0.60499	-1.01858	0.50920
				R negr	0.50920	0.49080	0.58313	0.98176	0.49080
				V bran	0.65644	0.34356	-0.42969	-0.72344	0.34356
				V negr	0.34356	0.65644	0.82102	1.38229	0.65644
Réu x Veredicto	0.02606	0.00068	0.22145 (1g1)	R bran	0.49080	0.50920	0.02655	1.01858	0.50920
				R negr	0.50920	0.49080	-0.02559	-0.98176	0.49080
				Condenado	0.11043	0.88957	0.07397	2.83823	0.88957
				Não conde	0.88957	0.11043	-0.00918	-0.35233	0.11043
Vítima x Veredicto	0.13124	0.01722	5.61487 (1g1)	V bran	0.65644	0.34356	0.09494	0.72344	0.34356
				V negr	0.34356	0.65644	-0.18141	-1.38229	0.65644
				Condenado	0.11043	0.88957	0.37249	2.83823	0.88957
				Não conde	0.88957	0.11043	-0.04624	-0.35233	0.11043

Gráfico 4.2.4 : Tabela Marginal Réu e Vítima

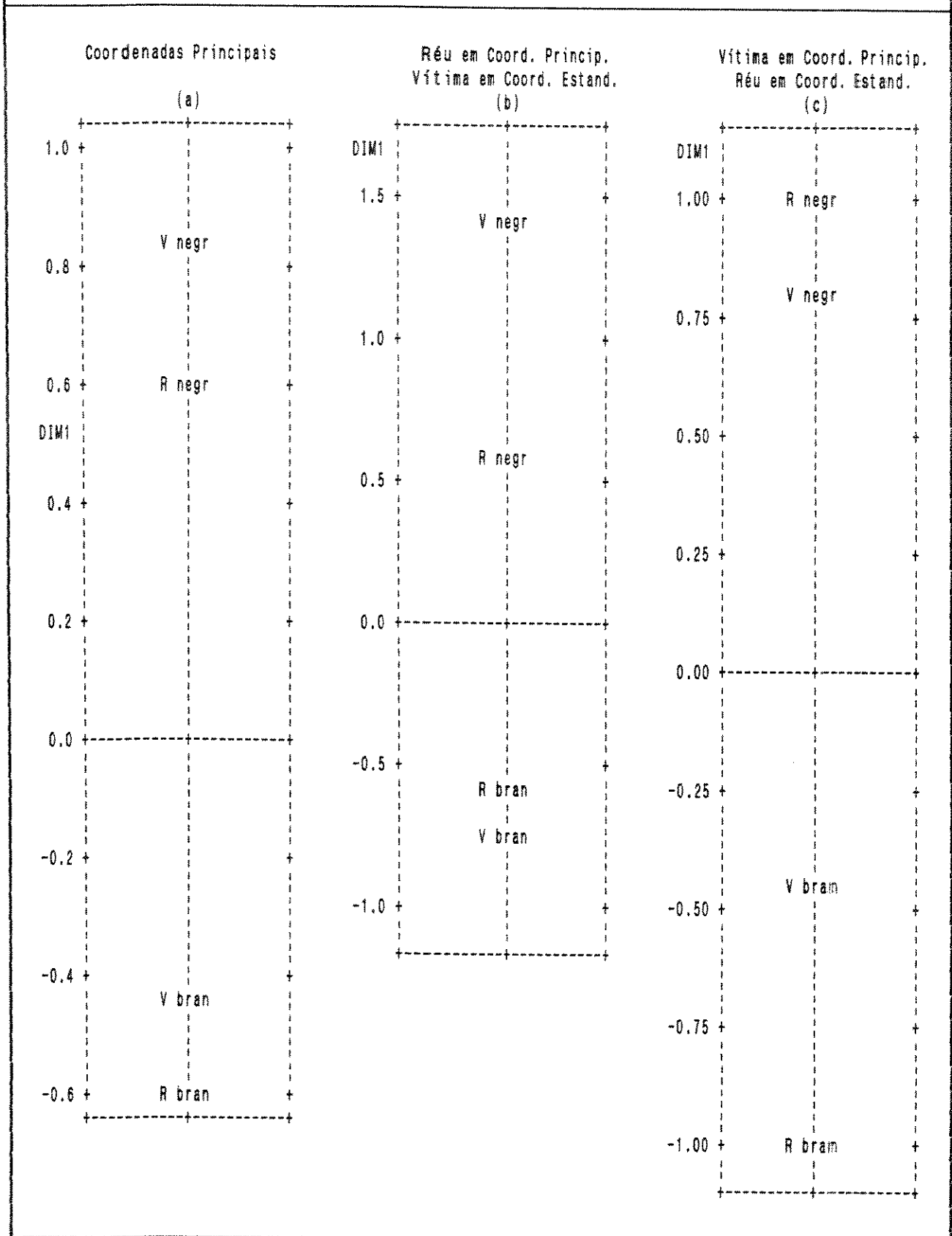


Gráfico 4.2.5 : Tabela Marginal Réu e Veredicto

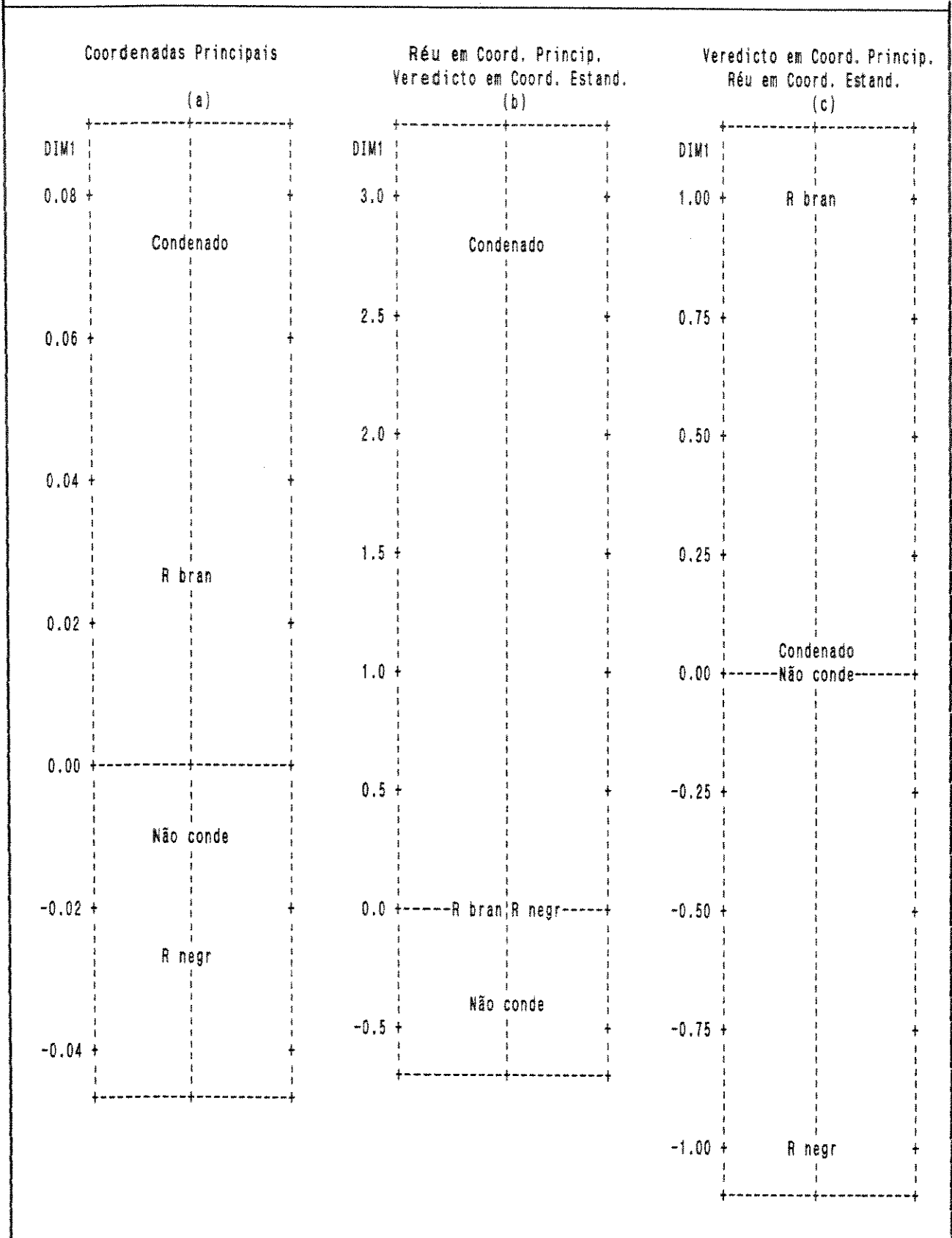
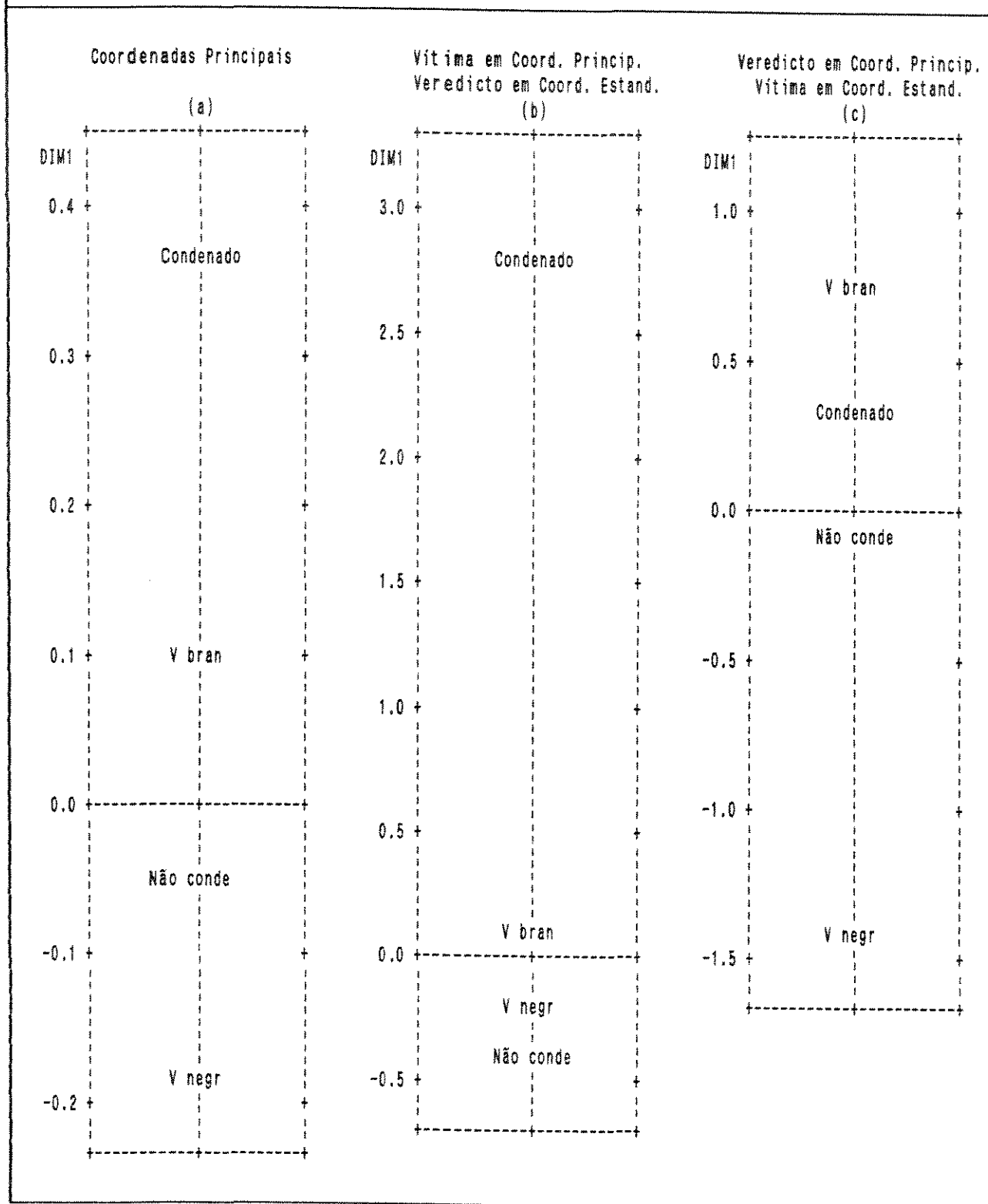


Gráfico 4.2.6 : Tabela Marginal Vítima e Veredicto



4.3 Aplicação da AC e ACG a um Exemplo de Etologia (Slater e Ollason, 1972)

Refaremos a análise de uma matriz de transição, que van der Heijden (1986) tratou com AC e ACG sob os modelos de independência e quase-independência respectivamente. Os dados foram tomados de Slater e Ollason (1972), destes temos as seguintes informações: Observou-se o comportamento de um grupo de pássaros machos "zebra finches", em cativeiro e isolados. Se anotaram os cambios de atividade para cada pássaro. As atividades consideradas para os pássaros foram:

Esticar, Locomover, Cantar, Enfeitar-se (se alisando com o bico),
Coçar-se, Esfregar (o bico), *Tomar-Banho* (de areia), *Posar,*
Beber, Comer, Ficar-Quieto, Arrepiar-se e Amolar (o bico)

A tabela de transição entre atividades, tem a propriedade particular que ela tem zeros estruturais fora da diagonal, desde que entre as quatro atividades *Banho, Beber, Comer e Amolar* não pode haver transição, pois estas atividades são feitas em locais diferentes e pelo menos entre elas deve haver *Locomover*.

A tabela 4.3.1 mostra as transições do pássaro nº 27. As caselas diagonais são zeros estruturais, e também as transições entre as atividades *Banho, Beber, Comer e Amolar*. As atividades no momento t estão representadas em minúsculas e as atividades no momento $t+1$ estão representadas em maiúsculas.

Tabela 4.3.1 : Transições entre as atividades do Pássaro nº 27 ("-" denota zeros estruturais).													
Atividades no momento <i>t</i>	Atividades no momento <i>t</i> +1												
	A	B	C	D	E	F	G	H	I	J	K	L	M
a esticar	-	3	3	3	0	2	0	0	1	0	1	1	0
b locomover	5	-	129	8	3	16	2	23	82	42	7	5	32
c cantar	2	121	-	21	2	3	0	4	1	3	2	11	4
d enfeitar	1	19	18	-	9	0	0	1	2	1	2	15	2
e coçar	0	2	4	7	-	1	0	0	0	0	1	1	0
f esfregar	2	25	2	6	0	-	0	0	6	2	0	2	0
g banho	0	2	0	0	0	0	-	0	-	-	0	0	-
h posar	0	18	2	0	1	1	0	-	3	3	1	2	2
i beber	3	69	0	3	1	16	-	1	-	-	2	0	-
j comer	1	39	1	5	0	2	-	1	-	-	3	1	-
k quieto	0	7	0	5	0	3	0	2	0	1	-	1	0
l arrepiar	0	16	12	8	2	1	0	0	0	0	0	-	0
m amolar	0	33	3	2	0	0	-	1	-	-	0	0	-

Fonte: van der Heijden (1986); Slater e Ollanson (1972)

4.3.1 Analisando a Tabela com AC para a decomposição dos resíduos do modelo de Independência

Primeiro, foi feita uma análise para investigar se existe alguma estrutura partindo da independência, sem considerar os zeros estruturais como tais.

A tabela 4.3.2 mostra a decomposição da inércia e da estatística Chi-quadrado, onde $\chi^2 = 898$, com 144 graus de liberdade, portanto, as frequências diferem de forma significativa do modelo de independência. Os três primeiros eixos contabilizam 89% da inércia.

As maiores contribuições ao primeiro eixo, estão dadas por *locomover*, com 56% no momento *t* e 44% no momento *t*+1. Do lado positivo do eixo, estão bem representados: *locomover* (*t*), junto com *banho*, *beber*, *comer*, *amolar* e *cantar* (*t*+1). Do lado negativo do eixo, estão bem representados *locomover* (*t*+1), junto com *banho*, *beber*, *comer*, *amolar* e *cantar* (*t*). Isto reflete a necessidade do pássaro se locomover para realizar tais atividades. Inclusive, podemos ver que o pássaro não faz nenhuma outra atividade no local onde toma *banho* de areia, pois antes e depois ele sempre se *locomove* (veja na tabela 4.3.3: as coordenadas principais de *banho* de areia (*t*) coincidem com as coordenadas standardizadas de *locomover* (*t*+1), isto significa que depois de

Tabela 4.3.3 : Estatísticas da AC para a Decomposição dos Resíduos do Modelo de Independência da Tabela de Transição entre as Atividades do Pássaro nº 27

Ativid mom t	Massa Prop. Inérc	Coordenad.Princip 1QEix 2QEix 3QEix	Coordenad.Estand. 1QEix 2QEix 3QEix	Contrib.Absoluta 1QEix 2QEix 3QEix	Contrib.Relativa 1QEix 2QEix 3QEix	Qual Repr
esticar	0.015 0.01	-0.07 -0.02 0.68	-0.09 -0.06 2.32	0.00 0.00 0.08	0.01 0.00 0.53	0.53
locomover	0.371 0.37	0.96 -0.12 -0.04	1.23 -0.32 -0.15	0.56 0.04 0.01	0.98 0.02 0.00	1.00
cantar	0.183 0.13	-0.78 0.02 0.03	-1.00 0.06 0.09	0.18 0.00 0.00	0.91 0.00 0.00	0.91
enfeitar	0.073 0.13	0.02 1.27 -0.14	0.03 3.37 -0.46	0.00 0.83 0.02	0.00 0.97 0.01	0.98
coçar	0.017 0.04	-0.15 0.04 1.53	-0.19 0.11 5.20	0.00 0.00 0.45	0.01 0.00 0.95	0.96
esfregar	0.047 0.02	-0.43 -0.11 0.07	-0.55 -0.30 0.24	0.01 0.00 0.00	0.38 0.03 0.01	0.42
banho	0.002 0.00	-1.09 -0.20 -0.49	-1.39 -0.53 -1.67	0.00 0.00 0.01	0.70 0.02 0.14	0.87
posar	0.035 0.01	-0.26 0.10 -0.38	-0.33 0.26 -1.29	0.00 0.00 0.06	0.19 0.03 0.42	0.64
beber	0.100 0.12	-0.83 -0.34 -0.29	-1.06 -0.91 -0.98	0.11 0.08 0.10	0.61 0.10 0.07	0.78
comer	0.056 0.05	-0.83 -0.16 -0.04	-1.06 -0.43 -0.12	0.06 0.01 0.00	0.85 0.03 0.00	0.88
quieto	0.020 0.03	-0.45 -0.29 0.66	-0.58 -0.77 2.25	0.01 0.01 0.10	0.16 0.06 0.34	0.56
arrepia	0.041 0.03	-0.24 0.14 0.57	-0.30 0.37 1.95	0.00 0.00 0.16	0.08 0.03 0.46	0.56
amolar	0.041 0.04	-0.85 -0.18 -0.22	-1.08 -0.47 -0.74	0.05 0.01 0.02	0.71 0.03 0.05	0.79
Ativid mom t+1	Massa Prop. Inérc	Coordenad.Princip 1QEix 2QEix 3QEix	Coordenad.Estand. 1QEix 2QEix 3QEix	Contrib.Absoluta 1QEix 2QEix 3QEix	Contrib.Relativa 1QEix 2QEix 3QEix	Qual Repr
ESTICAR	0.015 0.01	-0.08 -0.13 -0.26	-0.10 -0.36 -0.88	0.00 0.00 0.01	0.01 0.04 0.13	0.18
LOCOMOVER	0.371 0.30	-0.85 -0.07 -0.14	-1.09 -0.20 -0.49	0.44 0.01 0.09	0.96 0.01 0.03	0.99
CANTAR	0.183 0.16	0.85 0.12 0.11	1.09 0.33 0.37	0.22 0.02 0.02	0.90 0.02 0.01	0.94
ENFEITAR	0.071 0.09	-0.47 -0.13 0.99	-0.60 -0.36 3.37	0.03 0.01 0.81	0.18 0.01 0.79	0.98
COÇAR	0.019 0.06	-0.00 1.64 -0.16	-0.00 4.37 -0.53	0.00 0.36 0.01	0.00 0.91 0.01	0.92
ESFREGAR	0.047 0.05	-0.11 -0.49 -0.02	-0.14 -1.30 -0.06	0.00 0.08 0.00	0.01 0.22 0.00	0.23
BANHO	0.002 0.00	1.23 -0.32 -0.15	1.57 -0.86 -0.52	0.01 0.00 0.00	0.90 0.06 0.01	0.97
POSAR	0.035 0.02	0.61 -0.22 -0.03	0.77 -0.58 -0.10	0.02 0.01 0.00	0.58 0.07 0.00	0.66
BEBER	0.100 0.12	1.01 -0.22 -0.14	1.29 -0.58 -0.48	0.16 0.03 0.02	0.89 0.04 0.02	0.95
COMER	0.055 0.05	0.89 -0.20 -0.15	1.13 -0.54 -0.51	0.07 0.01 0.01	0.85 0.05 0.02	0.92
QUIETO	0.041 0.08	-0.21 1.24 0.02	-0.26 3.30 0.07	0.00 0.45 0.00	0.02 0.88 0.00	0.90
ARREPIAR	0.020 0.01	0.04 0.09 0.11	0.05 0.25 0.37	0.00 0.00 0.00	0.00 0.02 0.02	0.04
AMOLAR	0.042 0.04	0.87 -0.07 -0.20	1.11 -0.19 -0.68	0.05 0.00 0.02	0.86 0.01 0.05	0.91

Gráfico 4.3.1 : Coordenadas Principais dos Eixos Primeiro e Segundo da AC para a decomposição dos Resíduos do Modelo de Independência da Tabela de Transição entre as Atividades do Pássaro nº 27

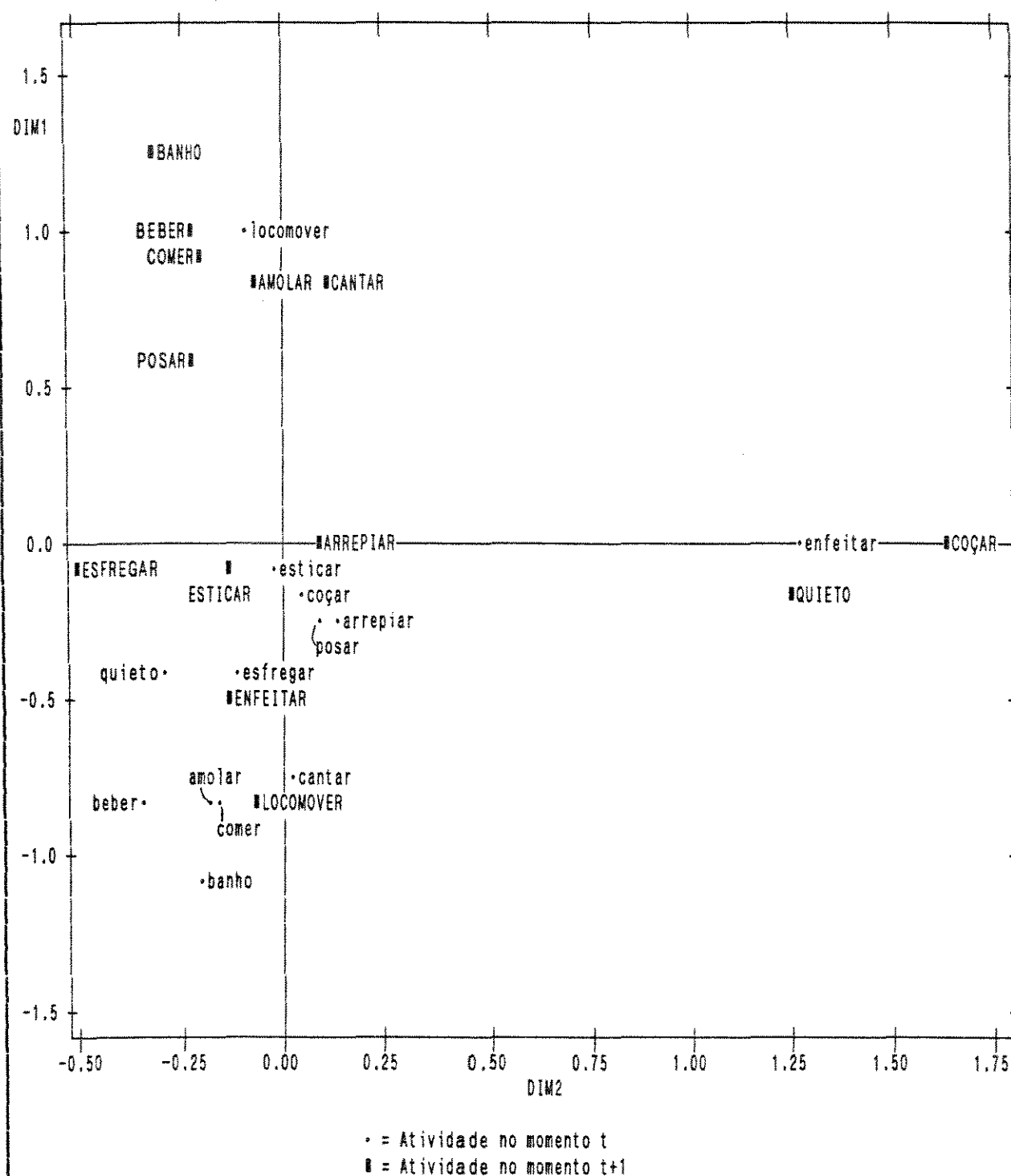
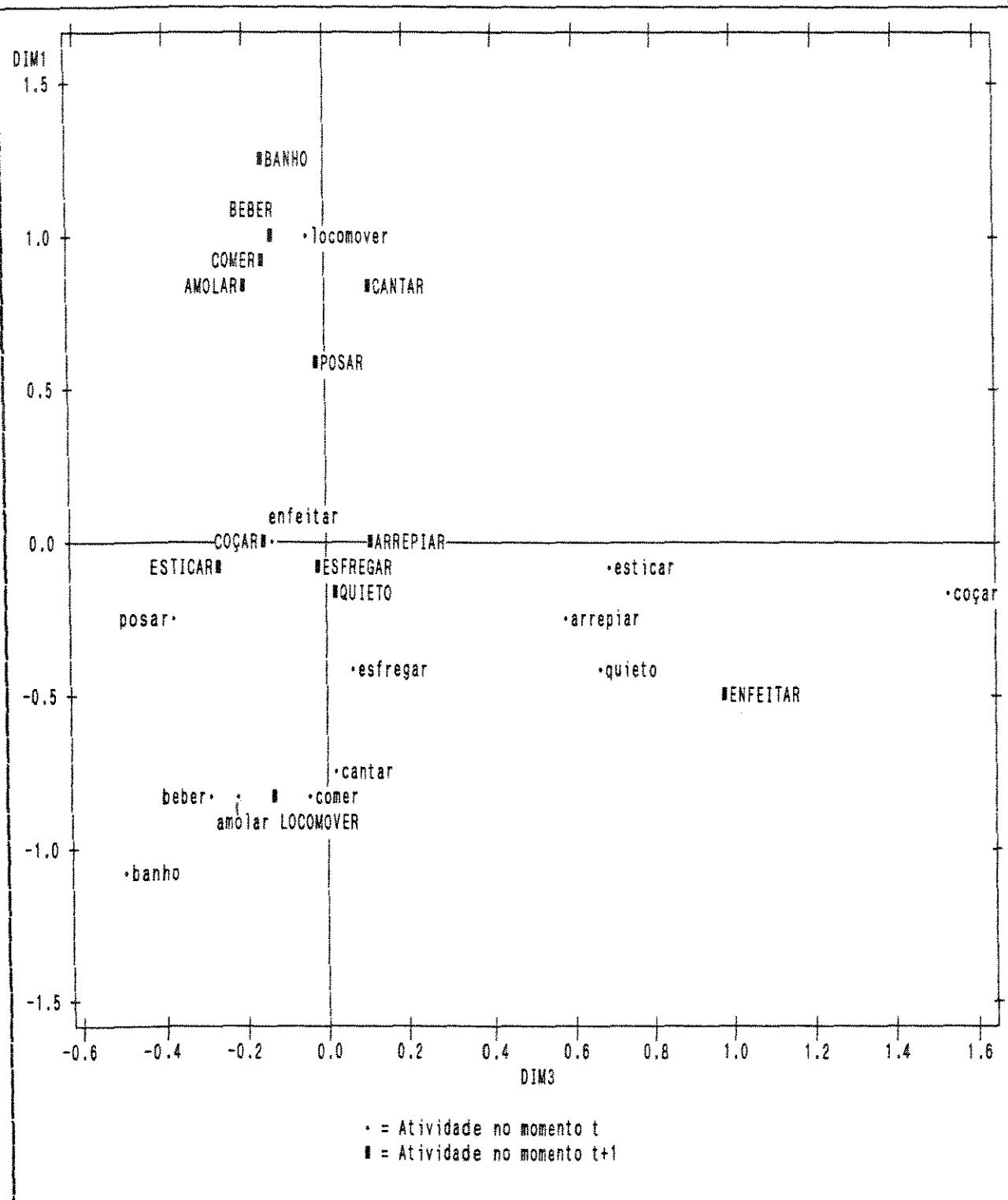


Gráfico 4.3.2 : Coordenadas Principais dos Eixos Primeiro e Terceiro da AC para a decomposição dos Resíduos do Modelo de Independência da Tabela de Transição entre as Atividades do Pássaro nº 27



4.3.2 Analisando a Tabela com ACG para a Decomposição dos Resíduos do Modelo de Quase-Independência

Uma solução que não é influenciada pelos zeros estruturais, pode ser encontrada usando a ACG, onde as diferenças entre as frequências observadas e as frequências esperadas sob o modelo de quase-independência são decompostas. Neste exemplo, para todos os zeros estruturais $q_{ij} = p_{ij} = 0$, e para as outras caselas $q_{ij} = a_i b_j$. Mas os cálculos foram feitos com o seguinte artifício: Primeiro, estimamos os valores a_i e b_j ($i, j = 1, \dots, 13$) da forma como está explicado na seção 3.5 (veja em anexo III.b). Logo, substituímos os valores $a_i b_j$ nas caselas com zeros estruturais. Daí, entramos com a tabela modificada no procedimento da AC. Desta maneira, a esperança sob independência da tabela modificada, apresentará os mesmos valores que a tabela modificada, conseqüentemente, estas não contribuem com a inércia, e temos a vantagem que as massas serão as apropriadas. Só não devemos esquecer o verdadeiro número de graus de liberdade que é igual a 119.

Os resíduos do modelo de quase-independência são também significativos, com o que a estatística $\chi^2 = 557$ com 119 graus de liberdade. A solução desta análise está nos gráficos 4.3.3 e 4.3.4. Os três primeiros eixos contabilizam 84% da inércia total.

Do lado positivo do primeiro eixo, podemos encontrar o comportamento de "arrumação" compreendidos pela atividades *coçar*, *quieto*, *enfeitar*, *arrepilar* e *esticar*. Do lado negativo do eixo, os comportamentos de "nutrição" (*beber*, *comer*) perto de *locomover* (t) e *LOCOMOVER* ($t+1$). Assim, o primeiro eixo pode ser interpretado como o cambio entre classes de atividades.

O segundo eixo, coloca em oposição as atividades no momento t e $t+1$.

Tabela 4.3.4 : Tabela de Decomposição da Inércia e a Estatística Chi-Quadrado por ACG, para a Decomposição dos Resíduos do Modelo de Quase-Independência da Tabela de Transição entre as Atividades do Pássaro nº 27

Eixo	Valores Singul.	Inércias Princip.	Chi-Quadr.	Percent	9	18	27	36	45
					-----+-----+-----+-----+-----+				
1º	0.41962	0.17608	250.076	44.90%	*****				
2º	0.31906	0.10180	144.576	25.96%	*****				
3º	0.22915	0.05251	74.574	13.39%	*****				
4º	0.15918	0.02534	35.988	6.46%	****				
5º	0.11275	0.01271	18.054	3.24%	**				
6º	0.10241	0.01049	14.896	2.67%	*				
7º	0.08733	0.00763	10.831	1.94%	*				
8º	0.06411	0.00411	5.837	1.05%	*				
9º	0.02949	0.00087	1.235	0.22%					
10º	0.02216	0.00049	0.698	0.13%					
11º	0.01200	0.00014	0.205	0.04%					
12º	0.00108	0.00000	0.002	0.00%					
Total		0.39217	556.971	(Graus de Liberdade = 119)					

Tabela 4.3.5 : Estatísticas da ACG para a Decomposição dos Resíduos do Modelo de Quase-Independência da Tabela de Transição entre as Atividades do Pássaro n2 27

Ativid mom t	Massa Prop. Inérc	Coordenad.Princip 1QEix 2QEix 3QEix	Coordenad.Estand. 1QEix 2QEix 3QEix	Contrib.Absoluta 1QEix 2QEix 3QEix	Contrib.Relativa 1QEix 2QEix 3QEix	Qual Repr
esticar	0.010 0.03	0.66 -0.74 0.08	1.58 -2.30 0.35	0.02 0.05 0.00	0.29 0.35 0.00	0.65
locomover	0.527 0.12	-0.23 0.15 -0.08	-0.55 0.47 -0.35	0.16 0.11 0.06	0.61 0.26 0.07	0.94
cantar	0.143 0.07	0.28 -0.16 -0.13	-0.66 -0.49 -0.56	0.06 0.03 0.04	0.39 0.12 0.08	0.60
enfeitar	0.052 0.31	1.30 0.72 0.36	3.09 2.27 1.58	0.50 0.27 0.13	0.72 0.22 0.06	1.00
coçar	0.011 0.11	1.28 -1.17 -0.61	3.04 -3.65 -2.66	0.11 0.15 0.08	0.44 0.37 0.10	0.91
esfregar	0.033 0.04	0.09 -0.36 0.07	0.21 -1.13 0.30	0.00 0.04 0.00	0.02 0.31 0.01	0.33
banho	0.002 0.00	-0.42 0.02 0.07	-1.00 0.04 0.31	0.00 0.00 0.00	0.35 0.00 0.01	0.36
posar	0.024 0.02	-0.02 0.23 0.30	-0.04 0.71 1.31	0.00 0.01 0.04	0.00 0.20 0.36	0.56
beber	0.079 0.12	-0.29 -0.39 0.58	-0.68 -1.21 2.52	0.04 0.12 0.50	0.14 0.25 0.56	0.94
comer	0.044 0.03	-0.09 -0.29 0.13	-0.22 -0.90 0.56	0.00 0.04 0.01	0.03 0.32 0.06	0.42
quieto	0.014 0.07	0.47 -1.07 0.18	1.13 -3.36 0.75	0.02 0.15 0.01	0.12 0.59 0.02	0.72
arrepilar	0.028 0.07	0.71 -0.25 -0.44	1.70 -0.78 -1.92	0.08 0.02 0.10	0.54 0.07 0.20	0.80
amolar	0.033 0.02	-0.26 -0.04 -0.10	-0.63 -0.14 -0.45	0.01 0.00 0.01	0.34 0.01 0.05	0.40
Ativid mom t+1	Massa Prop. Inérc	Coordenad.Princip 1QEix 2QEix 3QEix	Coordenad.Estand. 1QEix 2QEix 3QEix	Contrib.Absoluta 1QEix 2QEix 3QEix	Contrib.Relativa 1QEix 2QEix 3QEix	Qual Repr
ESTICAR	0.010 0.02	0.00 -0.24 0.53	0.01 -0.76 2.31	0.00 0.01 0.05	0.00 0.08 0.36	0.44
LOCOMOVER	0.527 0.05	-0.14 -0.03 0.01	-0.33 -0.05 0.06	0.06 0.00 0.00	0.53 0.02 0.00	0.56
CANTAR	0.143 0.09	0.20 0.26 -0.31	0.47 0.83 -1.35	0.03 0.10 0.26	0.15 0.28 0.38	0.80
ENFEITAR	0.051 0.24	0.88 -0.96 -0.37	2.10 -3.01 -1.62	0.22 0.46 0.13	0.42 0.50 0.07	0.99
COÇAR	0.013 0.15	1.69 0.99 0.63	4.03 3.10 2.76	0.21 0.12 0.10	0.64 0.22 0.09	0.95
ESFREGAR	0.033 0.12	-0.14 -0.76 0.74	-0.34 -2.39 3.21	0.00 0.19 0.34	0.01 0.42 0.39	0.82
BANHO	0.002 0.00	-0.55 0.24 -0.08	-1.30 0.77 -0.35	0.00 0.00 0.00	0.60 0.13 0.01	0.73
POSAR	0.024 0.02	-0.18 0.08 -0.10	-0.43 0.24 -0.44	0.00 0.00 0.00	0.10 0.02 0.03	0.15
BEBER	0.079 0.06	-0.40 0.17 0.04	-0.95 0.54 0.16	0.07 0.02 0.00	0.55 0.10 0.00	0.65
COMER	0.043 0.02	-0.35 0.13 0.06	-0.84 0.41 0.25	0.03 0.01 0.00	0.60 0.08 0.02	0.70
QUIETO	0.028 0.19	1.46 0.48 0.40	3.48 1.49 1.73	0.34 0.06 0.08	0.81 0.09 0.06	0.96
ARREPIAR	0.014 0.03	0.34 -0.23 0.29	0.81 -0.72 1.25	0.01 0.01 0.02	0.14 0.06 0.10	0.29
AMOLAR	0.033 0.02	-0.27 0.26 0.05	-0.65 0.80 0.23	0.01 0.02 0.00	0.42 0.37 0.02	0.81

Gráfico 4.3.3 : Coordenadas Principais dos Eixos Primeiro e Segundo da ACG para a decomposição dos Resíduos do Modelo de Quase-Independência da Tabela de Transição entre as Atividades do Pássaro nº 27

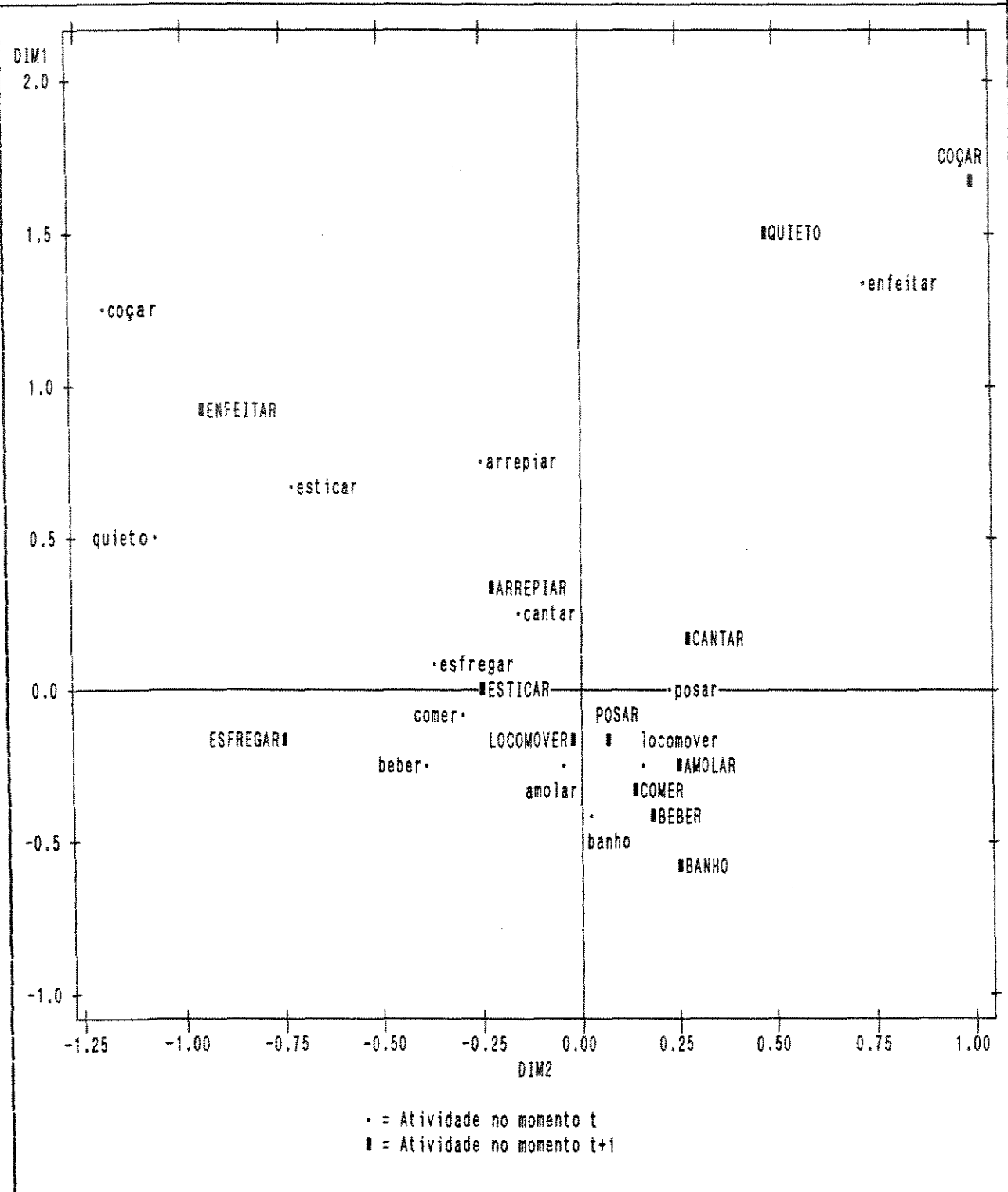
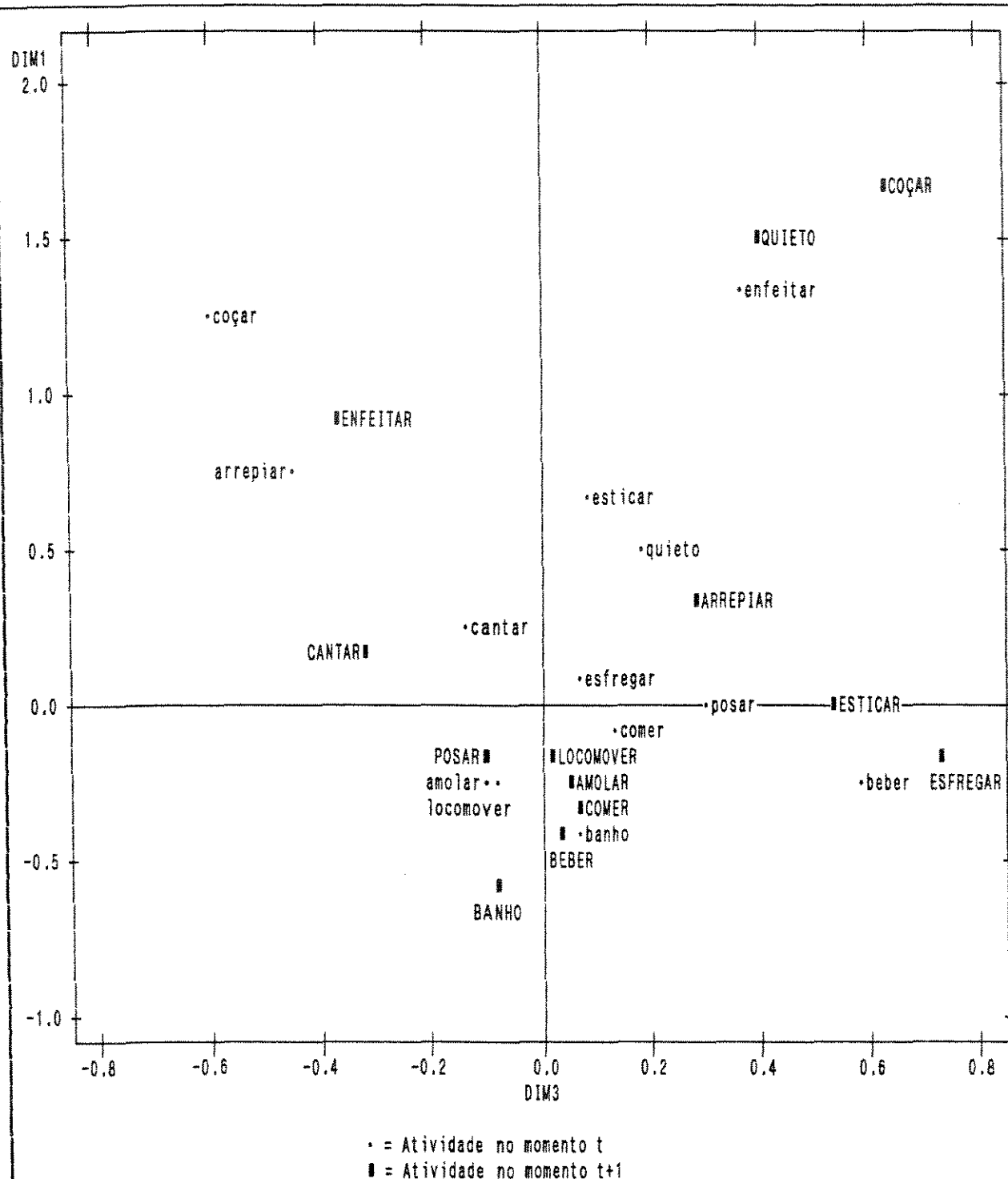


Gráfico 4.3.4 : Coordenadas Principais dos Eixos Primeiro e Terceiro da ACG para a decomposição dos Resíduos do Modelo de Quase-Independência da Tabela de Transição entre as Atividades do Pássaro nº 27



4.3.3 Analisando a Tabela com ACG para a Decomposição dos Resíduos do Modelo de Simetria

Para a aplicação da ACG a respeito do modelo de simetria, os cálculos foram feitos com o IML/SAS (veja anexo III.d).

A tabela 4.3.6 mostra que os dois primeiros eixos contabilizam 65% da inércia.

O gráfico 4.3.5 mostra que a configuração das atividades no momento t é igual que a configuração das atividades no momento $t+1$, mas com uma rotação de 90° no sentido anti-horário. Portanto, dos resultados de Constantine e Gower (1978), as preferências de transição entre atividades se dão no sentido horário, aproximadamente.

O gráfico 4.3.6 mostra somente as atividades no momento $t+1$ da tabela 4.3.1. Eliminando aquelas atividades com inércia muito pequena, menor que 0.03: *banho*, *amolar* e *cantar*, podemos observar que a transição entre atividades principalmente foi observada na seguinte seqüência fechada: *quieto*, *esfregar*, *enfeitar*, *arrepiar*, *locomover*, *posar*, *coçar*, *beber*, *comer*, *esticar*, e volta a ficar *quieto*...

Tabela 4.3.6 : Tabela de Decomposição da Inércia e a Estatística Chi-Quadrado por ACG, para a Decomposição dos Resíduos do Modelo de Simetria da Tabela de Transição entre as Atividades do Pássaro nº 27									
Eixo	Valores Singul.	Inércias Princip.	Chi-Quadr.	Percent	9	18	27	36	45
					-----+-----+-----+-----+-----+				
1º	0.15743	0.02478	23.6195	32.26%	*****				
2º	0.15743	0.02478	23.6195	32.26%	*****				
3º	0.08887	0.00790	7.5269	10.28%	*****				
4º	0.08887	0.00790	7.5269	10.28%	*****				
5º	0.06769	0.00458	4.3671	05.96%	**				
6º	0.06769	0.00458	4.3671	05.96%	**				
7º	0.03341	0.00112	1.0635	01.45%	*				
8º	0.03341	0.00112	1.0635	01.45%	*				
9º	0.00563	0.00003	0.0302	00.04%					
10º	0.00563	0.00003	0.0302	00.04%					
11º	0.00090	0.00000	0.0008	00.00%					
12º	0.00090	0.00000	0.0008	00.00%					
Total		0.07683	73.2160	(Graus de Liberdade = 78)					

Tabela 4.3.7 : Estatísticas da ACG para a Decomposição dos Resíduos do Modelo de Simetria da Tabela de Transição entre as Atividades do Pássaro nº 27

Ativid mom t	Massa	Porc. Inérc	Coord.Princ 1QEix 2QEix		Coord.Estan. 1QEix 2QEix		Contr.Absol 1QEix 2QEix		Contr.Relat 1QEix 2QEix		Qual Repr
esticar	0.015	0.05	-0.09	0.35	-0.55	2.20	0.00	0.07	0.03	0.46	0.49
locomover	0.371	0.09	0.06	-0.10	0.40	-0.67	0.06	0.16	0.23	0.57	0.80
cantar	0.183	0.02	0.03	0.04	0.24	0.28	0.01	0.01	0.17	0.23	0.40
enfeitar	0.072	0.18	-0.37	-0.15	-2.36	-0.93	0.40	0.06	0.73	0.11	0.83
coçar	0.018	0.04	0.26	-0.01	1.67	-0.08	0.05	0.00	0.37	0.00	0.37
esfregar	0.047	0.17	-0.39	0.23	-2.49	1.45	0.29	0.10	0.54	0.18	0.72
banho	0.002	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
posar	0.035	0.05	0.02	-0.03	0.16	-0.17	0.00	0.00	0.01	0.01	0.02
beber	0.100	0.12	0.17	0.18	1.11	1.17	0.12	0.14	0.34	0.38	0.72
comer	0.055	0.04	0.04	0.15	0.28	0.92	0.00	0.05	0.03	0.34	0.37
quieto	0.020	0.11	-0.20	0.51	-1.28	3.25	0.03	0.21	0.09	0.60	0.69
arrepiar	0.041	0.12	-0.10	-0.34	-0.66	-2.16	0.02	0.19	0.05	0.53	0.58
amolar	0.041	0.00	-0.01	0.00	-0.06	0.00	0.00	0.00	0.02	0.00	0.02

Ativid mom t+1	Massa	Prop. Inérc	Coord.Princ 1QEix 2QEix		Coord.Estan. 1QEix 2QEix		Contr.Absol 1QEix 2QEix		Contr.Relat 1QEix 2QEix		Qual Repr
ESTICAR	0.015	0.05	0.35	0.09	2.20	0.55	0.07	0.00	0.46	0.03	0.49
LOCOMOVER	0.371	0.09	-0.10	-0.06	-0.67	-0.40	0.16	0.06	0.57	0.21	0.78
CANTAR	0.183	0.02	0.04	-0.04	0.28	-0.24	0.01	0.01	0.23	0.17	0.80
ENFEITAR	0.072	0.18	-0.15	0.37	-0.93	2.36	0.06	0.40	0.11	0.73	0.84
COÇAR	0.018	0.04	-0.01	-0.26	-0.08	-1.67	0.00	0.05	0.00	0.37	0.37
ESFREGAR	0.047	0.17	0.23	0.39	1.45	2.49	0.10	0.29	0.18	0.54	0.72
BANHO	0.002	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
POSAR	0.035	0.05	-0.03	-0.03	-0.17	-0.16	0.00	0.00	0.01	0.01	0.02
BEBER	0.100	0.12	0.18	-0.17	1.17	-1.11	0.14	0.12	0.38	0.34	0.72
COMER	0.055	0.04	0.15	-0.04	0.92	-0.28	0.05	0.00	0.34	0.03	0.37
QUIETO	0.020	0.11	0.51	0.20	3.25	1.28	0.21	0.03	0.60	0.09	0.69
ARREPIAR	0.041	0.12	-0.34	0.10	-2.17	0.66	0.19	0.02	0.53	0.05	0.58
AMOLAR	0.041	0.00	0.00	0.01	0.00	0.06	0.00	0.00	0.00	0.02	0.02

Gráfico 4.3.5 : Coordenadas Principais dos Eixos Primeiro e Segundo da ACG para a decomposição dos Resíduos do Modelo de Simetria da Tabela de Transição entre as Atividades do Pássaro nº 27. Todos os pontos

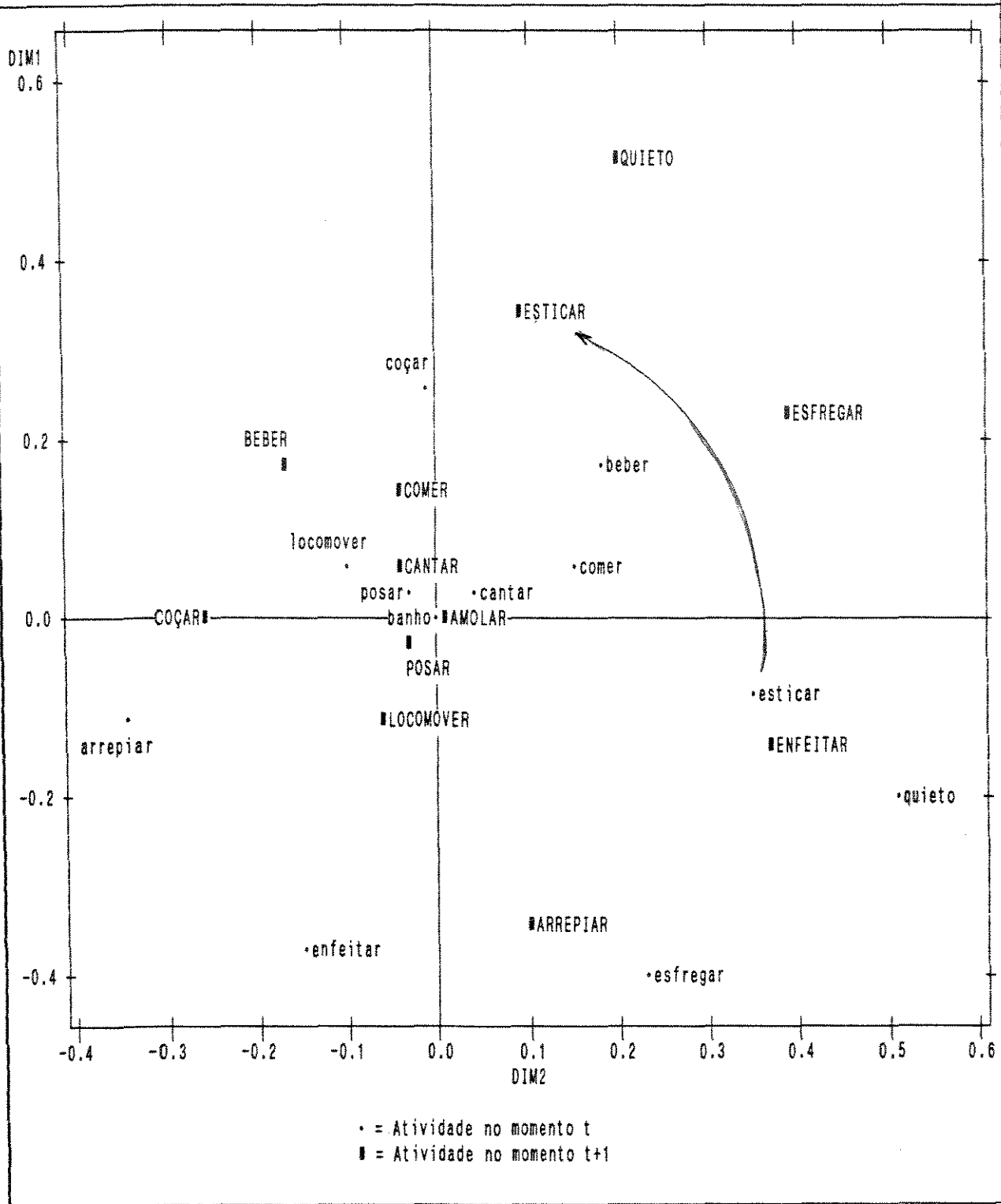
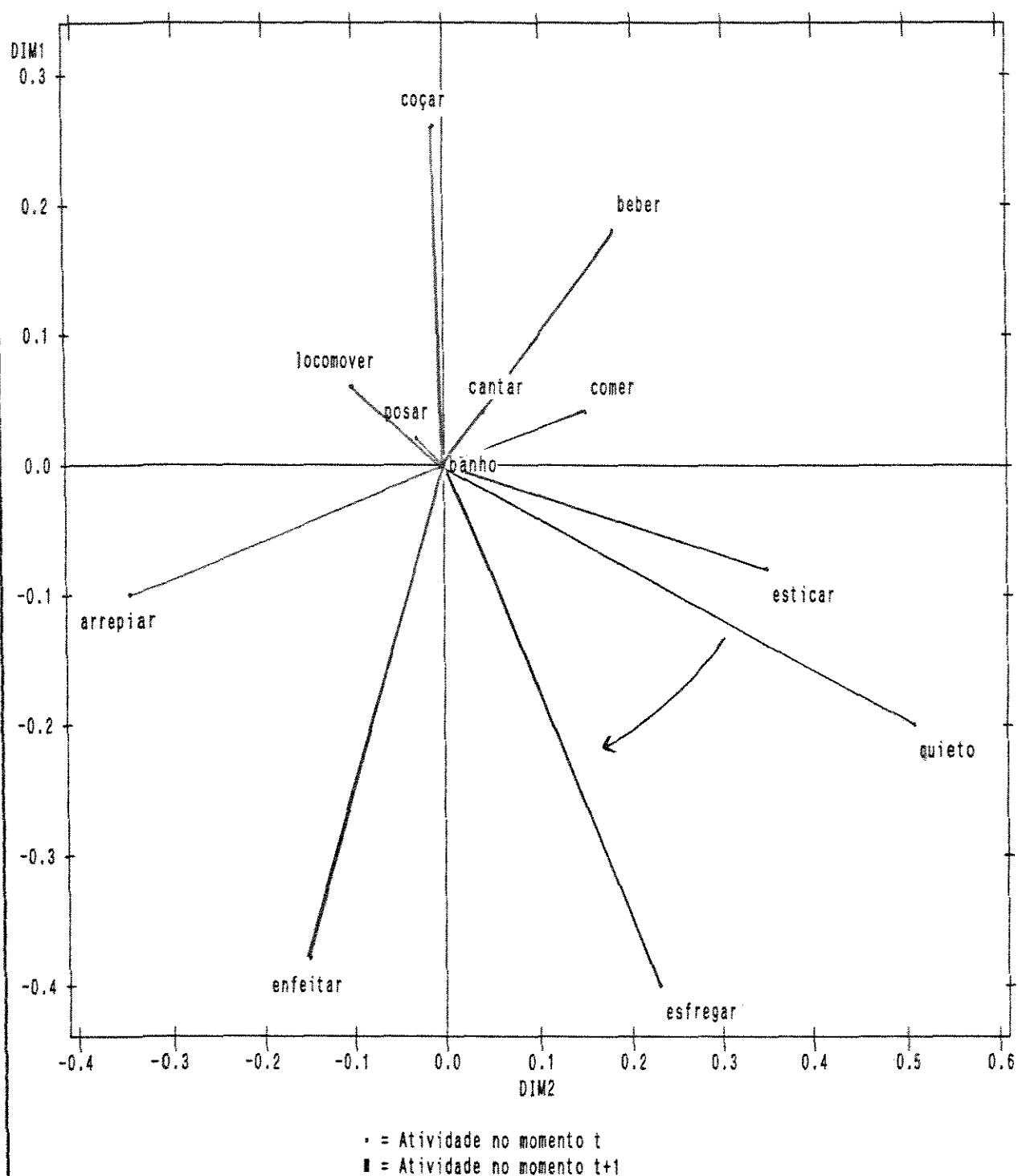


Gráfico 4.3.6 : Atividades(t) em Coordenadas Principais dos Eixos Primeiro e Segundo da ACG para a decomposição dos Resíduos do Modelo de Simetria da Tabela de Transição entre as Atividades do Pássaro nº 27



5. Discussão Geral e Conclusões

Enfocamos a interpretação da AC como análise de resíduos. Contudo, é importante usar AC nos casos onde os resíduos contêm informação significativa (por exemplo, quando o modelo Log-linear sob estudo não se ajusta bem). Mais ainda, AC faz sentido quando há algumas caselas cujas observações se distanciam de seus esperados.

O ganho na interpretação da AC como análise de resíduos, são primordialmente dois:

- 1º. Obtemos um melhor entendimento da AC relacionando-o à ALL.
- 2º. Tem-se demonstrado como a AC pode ser usada para a análise de resíduos de vários modelos log-lineares. A este respeito, AC pode auxiliar com a incômoda interpretação das interações ou dos resíduos quando temos um grande número de parâmetros. A ALL é usada para ajustar e testar modelos específicos; AC é usada para construir representações geométricas significativas de aspectos específicos dos dados ou dos modelos.

No contexto da AC, modelos log-lineares podem ter o seguinte papel (veja de Falguerolles e van der Heijden (1987)):

- 1º. Alguns modelos log-lineares podem ter um lugar central na análise, e AC pode ser usada para suplementar ALL com uma análise de resíduos destes modelos. De modo que, AC será útil se o distanciamento do modelo for significativo e se fizer sentido tentar estudar a estrutura dos resíduos.

29. O uso de modelos log-lineares permite eliminar a interação dominante (que pode ser desinteressante desde algum ponto de vista) das soluções da AC. Isto é útil se não estivermos interessados nos parâmetros mas só nos resíduos.
30. Como foi indicado na seção 3.3, AC pode ser usada como uma ferramenta exploratória para sugerir um modelo log-linear parsimonioso e facilmente interpretável a ser adotado como hipótese para estudos confirmatórios.

Como foi discutido na seção 3.3, os resultados de Goodman indicam a forma em que AC está relacionada com as abordagens baseadas em modelos, introduzindo os modelos de associação RC e de correlação RC (veja Goodman (1985, 1986)). As relações entre ACG e decomposições dos modelos ainda são pouco conhecidas, mas continuam como objeto de investigação.

Apêndices

A. A Distância Qui-Quadrado

O fato de trabalhar com distribuições condicionais nos conduz a utilizar uma distância distinta da euclidiana usual. A AC utiliza uma distância apropriada para distribuições, a distância *Qui-quadrado* que denotaremos por χ^2 .

- Distância χ^2 entre dois perfís linha i e l ($i, l = 1, \dots, I$):

$$d^2_{[R]}(i, l) = (\mathbf{r}_i - \mathbf{r}_l)' \mathbf{D}_c^{-1} (\mathbf{r}_i - \mathbf{r}_l) \quad (\text{A.1})$$

- Distância χ^2 entre dois perfís coluna j e k ($j, k = 1, \dots, J$):

$$d^2_{[C]}(j, k) = (\mathbf{c}_j - \mathbf{c}_k)' \mathbf{D}_r^{-1} (\mathbf{c}_j - \mathbf{c}_k). \quad (\text{A.2})$$

esta distância é uma distância euclidiana, mas difere da euclidiana usual unicamente na ponderação, sendo $\mathbf{D}_c^{-1}$ e $\mathbf{D}_r^{-1}$ nos espaços $\mathbb{R}^I$ e $\mathbb{R}^J$ respectivamente.

Utilizando a distância χ^2 estaremos dando um efeito estabilizador sobre os dados, visto que cumpre a propriedade denominada *Equivalência Distribucional*, que se expressa da seguinte maneira:

- Se duas linhas têm o mesmo perfil, elas podem ser substituídas por uma só, a qual tem o mesmo perfil, mas com uma massa igual à soma das massas. Sem que isto altere as distâncias entre os perfís coluna.

- Se duas colunas têm o mesmo perfil, elas podem ser substituídas por uma só, a qual tem o mesmo perfil, mas com uma massa igual a soma das massas. Sem que isto altere as distâncias entre os perfis linha.

Assim, se duas categorias com perfis iguais são juntadas, os resultados em geral não serão modificados.

Demonstração da Propriedade de Equivalência Distribucional

Nos limitaremos a demonstrar a propriedade para o cálculo da distância entre linhas, quando duas categorias coluna têm o mesmo perfil.

Suponhamos que as categorias coluna 1 e 2, são tais que têm o mesmo perfil, i.e.:

$$1/p_{.1} \mathbf{P}_1 = 1/p_{.2} \mathbf{P}_2 . \quad (\text{A.3})$$

onde para $j = 1, 2$: $1/p_{.j} \mathbf{P}_j = (p_{1j}/p_{.j}, p_{2j}/p_{.j}, \dots, p_{Ij}/p_{.j})'$.

Criaremos a categoria 0, como sendo uma nova categoria coluna que será a união das duas categorias 1 e 2, i.e.:

$$\mathbf{P}_0 = \mathbf{P}_1 + \mathbf{P}_2 , \quad (\text{A.4})$$

então a massa de 0 é igual a soma das massas, i.e.:

$$p_{.0} = p_{.1} + p_{.2} . \quad (\text{A.5})$$

Se as colunas 1 e 2 têm o mesmo perfil, não será difícil mostrar que o perfil da coluna 0 é igual aos perfis das categorias 1 e 2, i.e.:

$$1/p_{.0} \mathbf{P}_0 = 1/p_{.1} \mathbf{P}_1 = 1/p_{.2} \mathbf{P}_2 . \quad (\text{A.6})$$

Contando com as categorias coluna originais, podemos reescrever a distância entre dois perfis linha quaisquer i e l ($i, l = 1, \dots, I$) como:

$$d^2_{\{R\}}(i, l) = \sum_{j=1}^J T_j \quad (A.7)$$

onde $T_j = 1/p_{.j} (p_{ij}/p_{i.} - p_{lj}/p_{l.})^2$, para $j = 1, \dots, J$.

Se colocarmos a categoria 0 em vez das categorias originais 1 e 2, a distância entre os perfis linha se calcula segundo:

$$d^2_{\{R\}}(i, l) = T_0 + \sum_{j=3}^J T_j \quad (A.8)$$

onde $T_0 = 1/p_{.0} (p_{i0}/p_{i.} - p_{l0}/p_{l.})^2$.

Portanto, para mostrar que a propriedade é verdadeira, bastará verificar que se cumpre:

$$T_0 = T_1 + T_2. \quad (A.9)$$

Com efeito, para $n = 1, 2$ temos:

$$\begin{aligned} T_n &= \frac{1}{p_{.n}} \left[\frac{p_{in}}{p_{i.}} - \frac{p_{ln}}{p_{l.}} \right]^2 \\ &= p_{.n} \left[\frac{p_{in}}{p_{.n} p_{i.}} - \frac{p_{ln}}{p_{.n} p_{l.}} \right]^2 \end{aligned}$$

logo, por (A.6), temos:

$$T_n = p_{.n} \left[\frac{p_{i0}}{p_{.0} p_{i.}} - \frac{p_{l0}}{p_{.0} p_{l.}} \right]^2$$

assim:

$$\begin{aligned}
 T_1 + T_2 &= (p_{.1} + p_{.2}) \left[\frac{p_{i0}}{p_{.0} p_{i.}} - \frac{p_{10}}{p_{.0} p_{1.}} \right]^2 \\
 &= p_{.0} \left[\frac{p_{i0}}{p_{.0} p_{i.}} - \frac{p_{10}}{p_{.0} p_{1.}} \right]^2 \\
 &= \frac{1}{p_{.0}} \left[\frac{p_{i0}}{p_{i.}} - \frac{p_{10}}{p_{1.}} \right]^2 \\
 &= T_0 \blacksquare.
 \end{aligned}$$

Observações

- Quando juntamos duas colunas quaisquer, então as distâncias entre linhas diminui.
- Ao examinar a distância entre linhas, poderemos dizer que: se a diferença $(p_{ij}/p_{i.} - p_{1j}/p_{1.})$ for pequena com relação à massa da coluna j ($p_{.j}$), então, por meio da ponderação por $(1/p_{.j})$, estaremos evitando que esta diferença tenha muita influência na distância; por outro lado, se tal diferença for grande com relação à massa, então, com a ponderação, influenciará de forma proporcional na distância. Poderemos raciocinar de forma análoga quando observarmos a distância entre colunas.

B. Maximização de uma Forma Quadrática sob uma Restrição Quadrática

O Problema

Dada uma matriz simétrica $\mathbf{Z}$ de dimensão $J \times J$. Queremos encontrar um vetor $\mathbf{v}$ de $\mathbb{R}^J$, com $\mathbf{v}'\mathbf{v} = 1$, tal que:

$$\mathbf{v}' \mathbf{Z} \mathbf{v} = \max \{ \mathbf{x}' \mathbf{Z} \mathbf{x} : \mathbf{x} \in \mathbb{R}^J \text{ e } \mathbf{x}'\mathbf{x} = 1 \}.$$

Nota: No problema da AC, temos que a matriz $\mathbf{Z}$ é simétrica.

A Solução

Definimos a forma quadrática:

$$\begin{aligned} q : \mathbb{R}^J &\longrightarrow \mathbb{R} \\ \mathbf{x} &\longmapsto q(\mathbf{x}) = \mathbf{x}' \mathbf{Z} \mathbf{x} \end{aligned}$$

* Este é conhecido como Quociente de Rayleigh

Para encontrarmos a solução utilizaremos o método de *Multiplificadores de Lagrange*:

A equação de Lagrange para nosso problema é:

$$L(\mathbf{x}, \ell) = \mathbf{x}' \mathbf{Z} \mathbf{x} - \ell (\mathbf{x}' \mathbf{x} - 1) , \quad (\text{B.1})$$

onde ℓ é o multiplicador de Lagrange.

Para achar a solução optima, devemos igualar as derivadas da equação a zero. Busquemos a derivada da equação:

a) O gradiente da forma quadrática $q(\mathbf{x})$ é dado por::

$$\text{grad}[q(\mathbf{x})] = 2 \mathbf{Z} \mathbf{x} . \quad (\text{B.2})$$

Com efeito, sabemos que $\mathbf{x}' \mathbf{Z} \mathbf{x} = \sum_{ij} z_{ij} x_i x_j$, então para $k \in \{1, \dots, J\}$, temos:

$$\begin{aligned} \frac{d[\mathbf{x}' \mathbf{Z} \mathbf{x}]}{dx_k} &= \sum_{ij} z_{ij} \frac{d[x_i x_j]}{dx_k} \\ &= \sum_{ij} z_{ij} \left[x_j \frac{dx_i}{dx_k} + x_i \frac{dx_j}{dx_k} \right] \\ &= \sum_i \left(\sum_j z_{ij} x_j \right) \frac{dx_i}{dx_k} + \sum_j \left(\sum_i z_{ij} x_i \right) \frac{dx_j}{dx_k} , \\ &= \sum_j z_{kj} x_j + \sum_i z_{ik} x_i \\ &= 2 \sum_j z_{kj} x_j \\ &= 2 \mathbf{Z}_k' \mathbf{x} . \blacksquare \end{aligned}$$

b) Analogamente, temos que o gradiente de $\mathbf{x}'\mathbf{x}$ é:

$$\text{grad}[\mathbf{x}'\mathbf{x}] = 2 \mathbf{x}. \quad (\text{B.3})$$

c) Portanto, o gradiente da equação de Lagrange está dada por:

$$\text{grad}[\mathcal{L}(\mathbf{x}, \ell)] = 2 \mathbf{Z} \mathbf{x} - 2 \ell \mathbf{x} \quad (\text{B.4})$$

assim, obtemos o ponto crítico quando igualamos a derivada a zero ie:

$$\text{grad}[\mathcal{L}(\mathbf{v}, \ell)] = 2 \mathbf{Z} \mathbf{v} - 2 \ell \mathbf{v} = 0. \quad (\text{B.5})$$

Daqui, deduzimos a equação:

$$\mathbf{Z} \mathbf{v} = \ell \mathbf{v} \quad (\text{B.6})$$

Portanto temos que $\mathbf{v}$ é autovetor da matriz $\mathbf{Z}$. Quando pré-multiplicarmos a equação por $\mathbf{v}'$, e tomando em conta que $\mathbf{v}'\mathbf{v} = 1$, obtemos:

$$\ell = \mathbf{v}' \mathbf{Z} \mathbf{v} \quad (\text{B.7})$$

portanto, o valor do parâmetro ℓ será o maior autovalor de $\mathbf{Z}$.

Doravante, chamaremos $\mathbf{v}_1$ ao vetor $\mathbf{v}$ que corresponde ao maior valor ℓ_1 , tal que se cumpre a equação (B.6). Agora, procuraremos o vetor $\mathbf{v}_2$, que maximize a forma quadrática $\mathbf{v}_2' \mathbf{Z} \mathbf{v}_2$, e tal que seja ortogonal a $\mathbf{v}_1$ ($\mathbf{v}_2' \mathbf{v}_1 = 0$) e unitário ($\mathbf{v}_2' \mathbf{v}_2 = 1$). Análogamente, aplicaremos o método de Lagrange, obtendo a equação de Lagrange em:

$$\mathcal{L}_2(\mathbf{x}, \ell_2, m_2) = \mathbf{x}' \mathbf{Z} \mathbf{x} - \ell_2 (\mathbf{x}'\mathbf{x} - 1) - m_2 (\mathbf{x}'\mathbf{v}_1) \quad (\text{B.8})$$

onde ℓ_2 e m_2 são os multiplicadores de Lagrange.

O ótimo é encontrado quando igualamos a zero o gradiente da equação acima:

$$\text{grad}[\mathcal{L}_2(\mathbf{v}_2, \ell_2, m_2)] = 2 \mathbf{Z} \mathbf{v}_2 - 2 \ell_2 \mathbf{v}_2 - m_2 \mathbf{v}_1 = 0. \quad (\text{B.9})$$

Se pré-multiplicarmos esta equação por $\mathbf{v}_1'$, veremos que $m_2 = 0$, pois ao observar a transposta da expressão em (B.6) e pós-multiplicá-la por $\mathbf{v}_2$, obtemos:

$$\mathbf{v}_1' \mathbf{Z} \mathbf{v}_2 = \ell_1 \mathbf{v}_1' \mathbf{v}_2 = 0. \quad (\text{B.10})$$

Isto nos leva a:

$$\mathbf{Z} \mathbf{v}_2 = \ell_2 \mathbf{v}_2, \quad (\text{B.11})$$

assim, $\mathbf{v}_2$ é o segundo autovetor de $\mathbf{Z}$ associado ao segundo maior autovalor ℓ_2 , se este é único. Esta prova é facilmente extensível ao caso de o α -ésimo vetor $\mathbf{v}_\alpha$ ($\alpha = 1, \dots, J$), que maximize a forma quadrática $\mathbf{v}_\alpha' \mathbf{Z} \mathbf{v}_\alpha$, com $\mathbf{v}_\alpha' \mathbf{v}_\alpha = 1$ e $\mathbf{v}_\alpha' \mathbf{v}_{\alpha-1} = 0, \dots, \mathbf{v}_\alpha' \mathbf{v}_1 = 0$. Então obteremos:

$$\mathbf{Z} \mathbf{v}_\alpha = \ell_\alpha \mathbf{v}_\alpha. \quad (\text{B.12})$$

Enfim, temos que a solução ao problema está nos autovalores e autovetores de $\mathbf{Z}$:

$$\mathbf{Z} \mathbf{V} = \mathbf{V} \mathbf{D}_\ell, \quad (\text{B.13})$$

onde $\mathbf{V} = [\mathbf{v}_1 | \mathbf{v}_2 | \dots | \mathbf{v}_J]$ e $\mathbf{D}_\ell = \text{diag}\{\boldsymbol{\ell}\}$, com $\boldsymbol{\ell} = (\ell_1, \dots, \ell_J)'$, e $\ell_1 \geq \ell_2 \geq \dots \geq \ell_J$.

C. Decomposição em Valores Singulares

Se $\mathbf{X} = \mathbf{U} \mathbf{D}_{\mu} \mathbf{V}'$ com $\mathbf{U}'\mathbf{U} = \mathbf{I}$ e $\mathbf{V}'\mathbf{V} = \mathbf{I}$ então:

$$\begin{aligned}\mathbf{X}'\mathbf{X} &= \mathbf{V} \mathbf{D}_{\mu}^2 \mathbf{V}' \\ \mathbf{X} \mathbf{X}' &= \mathbf{U} \mathbf{D}_{\mu}^2 \mathbf{U}'.\end{aligned}\tag{C.1}$$

Reciprocamente, se uma matriz $\mathbf{X} = [x_{ij}]_{I \times J}$ cumpre:

$$\mathbf{X}' \mathbf{X} \mathbf{V} = \mathbf{V} \mathbf{D}_{\mu}^2 \tag{C.2}$$

onde:

$$\begin{aligned}\mathbf{V} &= [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_J], & \text{com } \mathbf{V}'\mathbf{V} &= \mathbf{I}_J \\ \mathbf{D}_{\mu} &= \text{diag}\{\boldsymbol{\mu}\}, & \text{com } \boldsymbol{\mu} &= (\mu_1, \dots, \mu_J)' \\ & & \text{e } \mu_1 &\geq \dots \geq \mu_J > 0,\end{aligned}$$

e

$$\mathbf{X} \mathbf{X}' \mathbf{U} = \mathbf{U} \mathbf{D}_{\tau}^2, \tag{C.3}$$

onde:

$$\begin{aligned}\mathbf{U} &= [\mathbf{u}_1 \mid \dots \mid \mathbf{u}_I], & \text{com } \mathbf{U}'\mathbf{U} &= \mathbf{I}_I \\ \mathbf{D}_{\tau} &= \text{diag}\{\boldsymbol{\tau}\}, & \text{com } \boldsymbol{\tau} &= (\tau_1, \dots, \tau_I)' \\ & & \text{e } \tau_1 &\geq \dots \geq \tau_I > 0.\end{aligned}$$

Então, existe um $D \leq \min\{I, J\}$, tal que:

- a. $\mu_{D+1} = \dots = \mu_J = \tau_{D+1} = \dots = \tau_I = 0$
- b. $\mathbf{D}_{\boldsymbol{\mu}^{(D)}} = \mathbf{D}_{\boldsymbol{\tau}^{(D)}}$
- c. $\mathbf{X} = \mathbf{U}^{(D)} \mathbf{D}_{\boldsymbol{\mu}^{(D)}} \mathbf{V}^{(D)}$,

onde:

$$\mathbf{U}^{(D)} = [\mathbf{u}_1 \mid \dots \mid \mathbf{u}_D],$$

$$\mathbf{V}^{(D)} = [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_D],$$

$$\mathbf{D}_{\boldsymbol{\mu}^{(D)}} = \text{diag}\{\boldsymbol{\mu}^{(D)}\}, \quad \text{com } \boldsymbol{\mu}^{(D)} = (\mu_1, \dots, \mu_D)', \text{ e}$$

$$\mathbf{D}_{\boldsymbol{\tau}^{(D)}} = \text{diag}\{\boldsymbol{\tau}^{(D)}\}, \quad \text{com } \boldsymbol{\tau}^{(D)} = (\tau_1, \dots, \tau_D)' .$$

Com efeito, se pré-multiplicarmos a equação (C.2) por $\mathbf{X}$, obteremos:

$$\mathbf{X} \mathbf{X}' (\mathbf{X} \mathbf{V}) = (\mathbf{X} \mathbf{V}) \mathbf{D}_{\boldsymbol{\mu}}^2$$

daqui, temos que, as colunas da matriz $(\mathbf{X} \mathbf{V})$ são autovetores de $\mathbf{X} \mathbf{X}'$ associados aos autovalores correspondentes em $\mathbf{D}_{\boldsymbol{\mu}}^2$. Se em $\boldsymbol{\mu}$ os valores são todos distintos, então a direção dos autovetores de $\mathbf{X} \mathbf{X}'$, são definidos de forma única; neste caso, teríamos que, cada coluna de $(\mathbf{X} \mathbf{V})$ é proporcional a coluna correspondente de $\mathbf{U}$, i.e. existe uma matriz diagonal $\mathbf{D}_{\mathbf{a}}$ tal que:

$$(\mathbf{X} \mathbf{V}) \mathbf{D}_{\mathbf{a}} = \mathbf{U} ,$$

logo, como $\mathbf{U}' \mathbf{U} = \mathbf{I}$, temos:

$$(\mathbf{X} \mathbf{V})' (\mathbf{X} \mathbf{V}) = \mathbf{D}_{\mathbf{a}}^{-2} ,$$

comparando este resultado com (C.2), temos que:

$$\mathbf{D}_{\mathbf{a}} = \mathbf{D}_{\boldsymbol{\mu}}^{-1} ,$$

assim:

$$\mathbf{U} = (\mathbf{X} \mathbf{V}) \mathbf{D}_{\boldsymbol{\mu}}^{-1}. \quad (\text{C.4})$$

Por outro lado, se pré-multiplicarmos a equação (C.3) por $\mathbf{X}'$, obteremos:

$$\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{U}) = (\mathbf{X}'\mathbf{U})\mathbf{D}_\tau^2$$

daqui, temos que, as colunas da matriz $(\mathbf{X}'\mathbf{U})$ são autovetores de $\mathbf{X}'\mathbf{X}$ associados aos autovalores correspondentes em $\mathbf{D}_\tau^2$. Se em τ os valores são todos distintos, então a direção dos autovetores de $\mathbf{X}'\mathbf{X}$, são definidos de forma única; neste caso, teríamos que, cada coluna de $(\mathbf{X}'\mathbf{U})$ é proporcional a coluna correspondente de $\mathbf{V}$, i.e. existe uma matriz diagonal $\mathbf{D}_b$ tal que:

$$(\mathbf{X}'\mathbf{U})\mathbf{D}_b = \mathbf{V} ,$$

logo, como $\mathbf{V}'\mathbf{V} = \mathbf{I}$, temos:

$$(\mathbf{X}'\mathbf{U})'(\mathbf{X}'\mathbf{U}) = \mathbf{D}_b^{-2} ,$$

comparando isto com (C.3), temos que:

$$\mathbf{D}_b = \mathbf{D}_\tau^{-1} ,$$

assim:

$$\mathbf{V} = (\mathbf{X}'\mathbf{U})\mathbf{D}_\tau^{-1}. \quad (\text{C.5})$$

De (C.4) e (C.5), obtemos:

$$\mathbf{X} = \mathbf{U}\mathbf{D}_\mu\mathbf{V}' \quad \text{e} \quad \mathbf{X}' = \mathbf{V}\mathbf{D}_\tau\mathbf{U}' , \quad (\text{C.6})$$

observemos que estes resultados são equivalentes. Observemos também que, só pode acontecer que, existe um $D \leq \min\{I, J\}$, tal que:

- a. $\mu_{D+1} = \dots = \mu_J = \tau_{D+1} = \dots = \tau_I = 0$.
- b. $\mathbf{D}_\mu^{(D)} = \mathbf{D}_\tau^{(D)}$.

onde:

$$\mathbf{D}_{\boldsymbol{\mu}}^{(D)} = \text{diag}\{\boldsymbol{\mu}^{(D)}\}, \quad \text{com } \boldsymbol{\mu}^{(D)} = (\mu_1, \dots, \mu_D)', \text{ e}$$

$$\mathbf{D}_{\boldsymbol{\tau}}^{(D)} = \text{diag}\{\boldsymbol{\tau}^{(D)}\}, \quad \text{com } \boldsymbol{\tau}^{(D)} = (\tau_1, \dots, \tau_D)' .$$

logo, refazendo o resultado em (C.6), obtemos:

$$\text{c. } \mathbf{X} = \mathbf{U}^{(D)} \mathbf{D}_{\boldsymbol{\mu}}^{(D)} \mathbf{V}^{(D)},$$

onde:

$$\mathbf{U}^{(D)} = [\mathbf{u}_1 \mid \dots \mid \mathbf{u}_D],$$

$$\mathbf{V}^{(D)} = [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_D].$$

Referências Bibliográficas

- AGRESTI, A. (1990) *Categorical Data Analysis*. New York, Wiley.
- ANDERSEN, E.B. (1980) *Discrete Statistical Models with Social Science Applications*. Amsterdam, North-Holland.
- BENZECRI, J.-P. (1973) *L'Analyse des Données*, vol. 1-2. Paris, Dunod.
- BISHOP, Y.M.M.; FIENBERG, S.E. AND HOLLAND, P.W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Massachusetts Institute of Technology Press.
- CABRAL, C.R.B. (1991) Modelos de Associação e Modelos de Correlação para Análise de Tabelas de Contingência com Duas Entradas. *Informes de Matemática*, Série D, Nº042. Rio de Janeiro, Instituto de Matemática Pura e Aplicada, fev/91.
- CALLEGARI-JAQUES, S.M. (1989) Análise de Correspondência: Aplicações em Genética. *Cadernos de Matemática e Estatística*, Série F, Nº2. Porto Alegre, Instituto de Matemática-UFRGS, dez/91.
- CAUSSINUS, H. (1965) Contribution à l'analyse statistique des tableaux de corrélation. *Ann. Fac. Sci. Univ. Toul.*, 29:77-182.
- CHOULAKIAN, V. (1988) Exploratory analysis of contingency tables by loglinear formulation and generalizations of correspondence analysis. *Psychometrika*, 53:235-250.
- CONSTANTINE, A.G. & GOWER, J.C. (1978) Graphical representation of asymmetric matrices. *Applied Statistics*, 27:297-304.
- DE LEEUW, J. & VAN DER HEIJDEN, P.G.M. (1988) Correspondence analysis of incomplete contingency tables. *Psychometrika*, 53:223-233.
- ECKART, C. & YOUNG, G. (1936) The approximation of one matrix by another of lower rank. *Psychometrika*, 1:211-218.

- ESCOFIER, B. (1983) Analyse de la différence entre deux mesures sur le produit de deux mêmes ensembles. *Cahiers de l'Analyse des Données*, 8(3):325-329.
- ESCOFIER, B. (1984) Analyse factorielle en référence à un modèle: application à l'analyse de tableaux d'échanges. *Revue de Statistique Appliquée*, 32(4):25-36.
- ESCOFIER, B. (1987) Using give for the computer program of correspondence analysis related to a model. *INRIA, Rapp. Rech.*, 616:1-13.
- ESCOUFIER, Y. (1985) L'analyse des correspondances: ses propriétés et ses extensions. *Bull. Int. Statist. Inst.*, 51(28.2):1-16.
- EVERITT, B.S. (1992) *The Analysis of Contingency Tables*. London, Chapman & Hall.
- FISHER, R.A. (1940) The precision of discriminant functions. *Ann. Eugen. Lond.*, 10:422-429.
- GABRIEL, K.R. (1971) The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3):453-467.
- GIFI, A. (1981) *Nonlinear Multivariate Analysis*. Leiden, The Netherlands, Department of Data Theory, University of Leiden.
- GILULA, Z. & HABERMAN, S.J. (1986) Canonical analysis of contingency tables by maximum likelihood. *Journal of the American Statistical Association*, 81:780-788.
- GOODMAN, L.A. (1979a) Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74:537-552.
- GOODMAN, L.A. (1979b) Multiplicative models for square contingency tables with ordered categories. *Biometrika*, 66:413-418.
- GOODMAN, L.A. (1981a) Association Models and Canonical Correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, 76:320-334.
- GOODMAN, L.A. (1981b) Association Models and the Bivariate Normal for Contingency Tables with Ordered Categories. *Biometrika*, 68:347-335.
- GOODMAN, L.A. (1985a) The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models and asymetry models for contingency tables with or without missing entries. *The Annals of Statistics*, 13(1):10-69.

- GOODMAN, L.A. (1985b) Correspondence analysis models, log-linear models, and log-bilinear models for the analysis of contingency tables. *Bull.Int.Statist.Inst.*, 51(28.1):14pp.
- GOODMAN, L.A. (1986) Some useful extentions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. *Int.Statist.Rev.*, 54:243-309.
- GOODMAN, L.A. (1991) Measures, Models and Graphical Displays in the Analysis of Cross-Classified Data. *Journal of the American Statistical Association*, 86(416):1085-1111,1124-1138.
- GOWER, J.C. (1977) The analysis of asymetry and orthogonality. In: *Recent Developments in Statistics* (ed.J.Barra et al.). Amsterdam, North-Holland.
- GRANDE ESTEBAN, I. & ABASCAL FERNANDEZ, E. (1989). *Métodos Multivariantes para la Investigación Comercial: teoría, aplicaciones y programación BASIC*. Barcelona, Ariel.
- GREENACRE, M.J. & HASTIE, T. (1987) The Geometric Interpretation of Correspondence Analysis. *Journal of the American Statistical Association*, 82(398):437-447.
- GREENACRE, M.J. (1984) *Theory and Applications of Correspondence Analysis*. New York, Academic Press.
- GUTTMAN, L. (1941) The quantification of a class of attributes: A theory and method of scale construction. In: *The Prediction of Personal Adjustment* (ed.Horst,P.). New York, Social Science Research Council, p.319-348.
- HILL, M.O. (1974) Correspondence Analysis: A Neglected Multivariate Method. *Applied Statistics*, 23(3):340-354.
- HIRSHFELD, H.O. (1935) A conection between correlation and contingency. *Cambridge Philos. Soc. Proc. (Math Proc.)*, 31:520-524.
- HORST, P. (1935) Measuring complex Attitudes. *Journal of Social Psychology*, 6:369-374.
- HUDON, G. (1990) A comparison of results of log-linear models and generalizations of correspondence analysis. *Revue de Statistique Appliquée.*, 38(2):43-53.
- LANCASTER, H.O. (1969) The Chi-Squared Distribution. New York, Wiley.
- LEBART, L. & FENELON, J.-P. (1973). *Statistique et Informatique Appliquées*. Paris, Dunod.

- LEBART, L.; MORINEAU, A. AND WARWICK, K.M. (1984). *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices*. New York, Wiley.
- MAQUERA SOSA, B.R. (1989) *Estudo de Tabelas de Contingência através da Análise de Correspondência*. Dissertação de Mestrado, IMECC/UNICAMP.
- MARDIA, K.V.; KENT, J.T. AND BIBBY, J.M. (1979) *Multivariate Analysis*. London, Academic Press.
- NISHISATO, S. (1980) *Analysis of Categorical Data: Dual Scaling and Its Applications*. Toronto, University of Toronto Press.
- NORA, C. (1975) Une méthode de reconstitution et d'analyse de données incomplètes. *Doctoral Dissertation*, Université Paul et Marie Curie, Paris.
- PEARSON, K. (1901) On Lines and Planes of Closest Fit to a System of points in Space. *Philosophical Magazine and Journal of Science, Series 6*, 2:559-572.
- RADELET, M. (1981) Racial Characteristics and the Imposition of the Death Penalty. *Amer. Sociol. Rev.* 46:918-927.
- RICHARDSON, M. & KUDER, G.F. (1933). Making a rating scale that measures. *Personnel Journal*, 12:36-40.
- SAS INSTITUTE INC. (1985) *SAS/IML User's Guide*, Version 5 Edition. Cary, NC: SAS Institute Inc..
- SAS INSTITUTE INC. (1989) *SAS/STAT User's Guide*, Version 6, Fourth Edition, Volume 1. Cary, NC: SAS Institute Inc..
- SLATER, P.J.B. & OLLASON J.C. (1972) The Temporal Patterning of Behavior Isolated Male Zebra Finches: Transition Analysis. *Behavior*, 42:248-269.
- SOUZA, A.M.R. DE (1982) *Análise de Correspondência*. Dissertação de Mestrado, Instituto de Matemática e Estatística, Universidade de São Paulo.
- SOUZA, N.M. DE (1990) *Análise de Correspondência*. Monografia, Bacharelado em Estatística, Universidade Federal do Rio Grande do Sul.
- TENEHAUS, M. & YOUNG, F. (1985) An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data. *Psychometrika*, 50:91-119.
- VAN DER HEIJDEN, P.G.M. & DE LEEUW, J. (1985) Correspondence Analysis used complementary to Log-linear Analysis. *Psychometrika*, 50(4):429-447.

- VAN DER HEIJDEN, P.G.M. & WORSLEY, K.J. (1988) Comment on Correspondence Analysis used complementary to log-linear Analysis (by P.G.M.van der Heijden and J.de Leeuw). *Psychometrika*, 53:287-291.
- VAN DER HEIJDEN, P.G.M. (1986) Transition Matrices, model fitting, and correspondence analysis. E.Diday In: *Data Analysis and Informatics IV*. Amsterdam, North-Holland.
- VAN DER HEIJDEN, P.G.M. (1987) *Correspondence Analysis of Logitudinal Categorical Data*. Leiden, DSWO.
- VAN DER HEIJDEN, P.G.M.; DE FALGUEROLLES, A. AND DE LEEUW, J. (1989) A combined approach to contingency table analysis using correspondence analysis and log-linear analysis (with discussion). *Applied Statistics*, 38(2):249-292.
- VERDINELLI, M.A. (1980) *Análise Inercial em Ecologia*. Tese de Doutorado, Instituto Oceanográfico, Universidade de São Paulo, São Paulo.
- WORSLEY, K.J. (1987) Un exemple d'identification d'un modèle log-linéaire grâce à une analyse des correspondances. *Revue de Statistique Appliquée*, 35(3):13-20, 83-94.

Anexos

I. Exemplo da Lingüística

I.a Listagem do Programa em SAS para AC

```
title 'Pesquisa da Lingüística';
title2 'Classificação das Palavras com Relação ao Teste de Familiaridade';

data LIdados;
input
    Palavra$ 1-14 TFam1 21-22 TFam2 29-30 TFam3 37-38
;
Label
    Palavra = 'Palavra'
    TFam1   = '1 Não Familiar'
    TFam2   = '2 Pouco Familiar'
    TFam3   = '3 Familiar'
;
cards;
A*loadings          45      3      1
B cluster           37      7      5
C*skapping          41      6      2
D spawned           45      4      0
E shaken            21      7     21
F*glim              47      2      0
G loomed            47      1      1
H*drealled          48      1      0
I resembling        40      8      1
J cloud             18     10     21
K*team              44      4      1
L disk-shaped       27     15      7
M*tark              47      0      1
N fortuitously      27     10     12
O fairly            21     17     11
P*observement       16     15     16
Q massive           26      7     16
R*trivel            45      3      0
S picked up         17     13     19
T focusing          34      9      6
U spring            9       5     35
V empty             6       4     39
W throughout        15     21     13
X pattern           17     14     18
Y appeared          14     10     25
Z close             3       2     44
a*adjustion         15     16     18
b recieving         6      12     31
c*resks             40      8      1
d slowly            6       1     42;
run;
```

```

title3 'Com as Coordenadas Principais';
proc corresp
    data = LIdados outc = LIout
    dimens = 2
    profile = both
    rp
    short;
    var TFam1 TFam2 TFam3;
    id Palavra;
    run;
proc print data = LIout;
    run;
proc plot
    data = LIout;
    plot dim1*dim2 = Palavra / box href = 0 vref = 0 ;
    run;

title3 'Palavras em Coordenadas Principais e Familiaridade em Coordenadas Estandardizadas';
proc corresp
    data = LIdados outc = LIout
    dimens = 2
    profile=row
    short;
    var TFam1 TFam2 TFam3;
    id Palavra;
    run;
proc print data = LIout;
    run;
proc plot
    data = LIout;
    plot dim1*dim2 = Palavra / box href = 0 vref = 0 ;
    run;

/*-----*/
/*          Fim de Programa          */
/*-----*/

```

II. Exemplo da Pena de Morte

II.a Listagem do Programa em SAS para ALL

```
title 'Veredito da Pena de Morte segundo a Raça do Réu e a Raça da Vítima';
title2 'Tabela de Três Entradas: Réu Vítima e Veredito';
```

```
data PMdados;
```

```
input
```

```
Reu$ 1-6 Vitima$ 8-13 PMorte$ 15-23 numero 25-27;
```

```
Label
```

```
Reu    = 'Raça do Réu'
```

```
Vitima = 'Raça da Vítima'
```

```
PMorte = 'Veredito de Pena de Morte'
```

```
;
```

```
cards;
```

```
R bran V bran Condenado 19
```

```
R bran V bran Não conde 132
```

```
R bran V negr Condenado 0
```

```
R bran V negr Não conde 9
```

```
R negr V bran Condenado 11
```

```
R negr V bran Não conde 52
```

```
R negr V negr Condenado 6
```

```
R negr V negr Não conde 97
```

```
;
```

```
run;
```

```
/*-----*/
/*          Tabela de Três Entradas          */
/*-----*/
```

```
title3 'Tabela de Três Entradas, Modelo [12][23]';
```

```
proc catmod data = PMdados
```

```
order = data;
```

```
weight numero;
```

```
model Reu*Vitima*PMorte = _response_ /
```

```
ml nodesign noiter noprofile noresponse;
```

```
loglin Reu Vitima PMorte
```

```
Reu*Vitima Vitima*PMorte;
```

```
run;
```



```

/*-----*/
/*          Tabelas Marginais          */
/*-----*/

```

```

title3 'Tabela Marginal Réu Vítima';
proc catmod data = PMdados
  order = data;
  weight numero;
  model Reu*Vítima = _response_ /
    ml nodelink noiter noprofile noresponse;
  loglin Reu Vítima Reu*Vítima;
run;

```

```

title3 'Tabela Marginal Réu Veredicto';
proc catmod data = PMdados
  order = data;
  weight numero;
  model Reu*PMorte = _response_ /
    ml nodelink noiter noprofile noresponse;
  loglin Reu PMorte;
run;

```

```

title3 'Tabela Marginal Vítima Veredicto';
proc catmod data = PMdados
  order = data;
  weight numero;
  model Vítima*PMorte = _response_ /
    ml nodelink noiter noprofile noresponse;
  loglin Vítima PMorte Vítima*PMorte;
run;

```

```

/*-----*/
/*          Fim de Programa          */
/*-----*/

```

II.b Listagem do Programa em SAS para AC

```
title 'Veredicto da Pena de Morte segundo a Raça do Réu e a Raça da Vítima';
```

```
data PMdados;
```

```
input
```

```
Reu$ 1-6 Vitima$ 8-13 PMorte$ 15-23 numero 25-27;
```

```
Label
```

```
Reu = 'Raça do Réu'
```

```
Vitima = 'Raça da Vítima'
```

```
PMorte = 'Veredicto de Pena de Morte'
```

```
;
```

```
cards;
```

```
R bran V bran Condenado 19
```

```
R bran V bran Não conde 132
```

```
R bran V negr Condenado 0
```

```
R bran V negr Não conde 9
```

```
R negr V bran Condenado 11
```

```
R negr V bran Não conde 52
```

```
R negr V negr Condenado 6
```

```
R negr V negr Não conde 97
```

```
;
```

```
run;
```

```
/*-----*/  
/*          Tabelas Múltiplas, Combinando duas Vars          */  
/*-----*/
```

```
title2 'Tabela Combinando Réu e Vítima contra Veredicto';
```

```
title3 'Coordenadas F e G';
```

```
proc corresp data = PMdados outc = PMout
```

```
cross = row
```

```
dimens = 1
```

```
profile = both
```

```
short;
```

```
tables Reu Vitima , Pmorte;
```

```
weight numero;
```

```
run ;
```

```
proc print data = PMout;
```

```
data PMout; set PMout; dim2 = 0;
```

```
proc plot data = PMout hpercent=40;
```

```
plot dim1*dim2 = _name_/box vref=0 href=0;
```

```
run;
```

```
title3 'Coordenadas F e I';
```

```
proc corresp data = PMdados outc = PMout
```

```
cross = row
```

```
dimens = 1
```

```
profile = row
```

```
short ;
```

```
tables Reu Vitima , Pmorte;
```

```

weight numero ;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent=40;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title2 'Tabela Combinando Réu e Veredicto contra Vítima';
title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
cross = row
dimens = 1
profile = both
short;
tables Reu Pmorte , Vítima;
weight numero;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent=40;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
cross = row
dimens = 1
profile = row
short ;
tables Reu Pmorte , Vítima;
weight numero ;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent=40;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title2 'Tabela Combinando Vítima e Veredicto contra Réu';
title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
cross = row
dimens = 1
profile = both
short;
tables Vítima Pmorte , Reu;
weight numero;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent=40;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title3 'Coordenadas F e G';

```

```

proc corresp data = PMdados outc = PMout
  cross = row
  dims = 1
  profile = row
  short ;
  tables Vitima Pmorte , Reu;
  weight numero ;
  run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent=40;
  plot dim1*dim2 = _name_/box vref=0 href=0;
  run;

/*-----*/
/*          Tabelas Marginais          */
/*-----*/

title2 'Tabela Marginal Vitima e Veredicto';
title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
  dims = 1
  profile = both
  short;
  tables Vitima, Pmorte;
  weight numero;
  run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
  plot dim1*dim2 = _name_/box vref=0 href=0;
  run;
title3 'Coordenadas F e T';
proc corresp data = PMdados outc = PMout
  dims = 1
  profile = row
  short;
  tables Vitima, Pmorte;
  weight numero;
  run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
  plot dim1*dim2 = _name_/box vref=0 href=0;
  run;
title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
  dims = 1
  profile = column
  short;
  tables Vitima, Pmorte;
  weight numero;
  run ;

```

```

proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
    plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title2 'Tabela Marginal Réu e Veredicto';
title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
    dims = 1
    profile = both
    short;
tables Reu, Pmorte;
weight numero;
run ;

proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
    plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title3 'Coordenadas F e R';
proc corresp data = PMdados outc = PMout
    dims = 1
    profile = row
    short;
tables Reu, Pmorte;
weight numero;
run ;

proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
    plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title3 'Coordenadas R e G';
proc corresp data = PMdados outc = PMout
    dims = 1
    profile = column
    short;
tables Reu, Pmorte;
weight numero;
run ;

proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
    plot dim1*dim2 = _name_/box vref=0 href=0;
run;

title2 'Tabela Marginal Réu e Vitima';
title3 'Coordenadas F e G';
proc corresp data = PMdados outc = PMout
    dims = 1
    profile = both
    short;
tables Reu, Vitima;

```

```

weight numero;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;
title3 'Coordenadas F e Γ';
proc corresp data = PMdados outc = PMout
dimens = 1
profile = row
short;
tables Reu, Vitima;
weight numero;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;
title3 'Coordenadas Φ e G';
proc corresp data = PMdados outc = PMout
dimens = 1
profile = column
short;
tables Reu, Vitima;
weight numero;
run ;
proc print data = PMout;
data PMout; set PMout; dim2 = 0;
proc plot data = PMout hpercent = 30;
plot dim1*dim2 = _name_/box vref=0 href=0;
run;

```

```

/*-----*/
/*          Fim de Programa          */
/*-----*/

```

III. Exemplo da Etologia

III.a Listagem do Programa em SAS para AC, aplicada para a Decomposição dos Resíduos do Modelo de Independência

```
title 'Etologia em Pássaros';
title2 'Decomposição dos Resíduos do Modelo de Independência';
title3 'Considerando os Zeros Estruturais como Zeros Aleatórios';

data EPdados;

input
  Ativid$ 6-18 A B C D E F G H I J L K M;

Label
  A = 'A ESTICAMENTO'
  B = 'B LOCOMOÇÃO '
  C = 'C CANTANDO '
  D = 'D ENFEITANDO '
  E = 'E COÇANDO '
  F = 'F ESFREG.BICO'
  G = 'G BANHO AREIA'
  H = 'H ESTEREOTIP '
  I = 'I BEBENDO '
  J = 'J COMENDO '
  K = 'K QUIETO '
  L = 'L ARREPIANDO '
  M = 'M AMOLA.BICO '
;

cards;
  a esticamento 0 3 3 3 0 2 0 0 1 0 1 1 0
  b locomoção 5 0 129 8 3 16 2 23 82 42 7 5 32
  c cantando 2 121 0 21 2 3 0 4 1 3 2 11 4
  d enfeitando 1 19 18 0 9 0 0 1 2 1 2 15 2
  e coçando 0 2 4 7 0 1 0 0 0 0 1 1 0
  f esfreg.bico 2 25 2 6 0 0 0 0 6 2 0 2 0
  g banho areia 0 2 0 0 0 0 0 0 0 0 0 0 0
  h estereotip 0 18 2 0 1 1 0 0 3 3 1 2 2
  i bebendo 3 69 0 3 1 16 0 1 0 0 2 0 0
  j comendo 1 39 1 5 0 2 0 1 0 0 3 1 0
  k quieto 0 7 0 5 0 3 0 2 0 1 0 1 0
  l arrepiando 0 16 12 8 2 1 0 0 0 0 0 0 0
  m amola.bico 0 33 3 2 0 0 0 1 0 0 0 0 0
;

run;

title4 'Coordenadas F e G';
proc corresp
  data = EPdados outc = EPout
```

```

    dimens = 3
    profile = both
    observed rp cp
    short;
    var A B C D E F G H I J L K M;
    id Ativid;
    run;
proc print data = EPout;
    run;
proc plot
    data = EPout;
    plot dim1*dim2 = Ativid / box href = 0 vref = 0 ;
    plot dim1*dim3 = Ativid / box href = 0 vref = 0 ;
    plot dim2*dim3 = Ativid / box href = 0 vref = 0 ;
    run;

title4 'Coordenadas  $\Phi$  e  $\Gamma$ ';
proc corresp
    data = EPdados outc = EPout
    dimens = 3
    row = DA column = DB
    short;
    var A B C D E F G H I J L K M;
    id Ativid;
    run;

/*-----*/
/*          Fim de Programa          */
/*-----*/

```


III.b Listagem da Rodada do Programa em IML para a Atribuição de Valores para as Caselas com Zeros Estruturais pelo Algoritmo EM

```

1  /*-----*/
2  /* Etologia em Pássaros */
3  /*-----*/
4
5  proc IML ;
IML Ready
6
7
8  * A Tabela de Contingência TC ;
9  TC = { 0 3 3 3 0 2 0 0 1 0 1 1 0 ,
10         5 0 129 8 3 16 2 23 82 42 7 5 32 ,
11         2 121 0 21 2 3 0 4 1 3 2 11 4 ,
12         1 19 18 0 9 0 0 1 2 1 2 15 2 ,
13         0 2 4 7 0 1 0 0 0 0 1 1 0 ,
14         2 25 2 6 0 0 0 0 6 2 0 2 0 ,
15         0 2 0 0 0 0 0 0 0 0 0 0 0 ,
16         0 18 2 0 1 1 0 0 3 3 1 2 2 ,
17         3 69 0 3 1 16 0 1 0 0 2 0 0 ,
18         1 39 1 5 0 2 0 1 0 0 3 1 0 ,
19         0 7 0 5 0 3 0 2 0 1 0 1 0 ,
20         0 16 12 8 2 1 0 0 0 0 0 0 0 ,
21         0 33 3 2 0 0 0 1 0 0 0 0 0 } ;
22
23
24  * Inicialização de Variáveis Auxiliares;
25  ui = { 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 1 } ;
26
27  Atr = TC ;
28  Atr[01,01] = 0 ; Atr[02,02] = 0 ; Atr[03,03] = 0 ; Atr[04,04] = 0 ;
29  Atr[05,05] = 0 ; Atr[06,06] = 0 ; Atr[07,07] = 0 ; Atr[08,08] = 0 ;
30  Atr[09,09] = 0 ; Atr[10,10] = 0 ; Atr[11,11] = 0 ; Atr[12,12] = 0 ;
31  Atr[13,13] = 0 ; Atr[07,09] = 0 ; Atr[07,10] = 0 ; Atr[07,13] = 0 ;
32  Atr[09,07] = 0 ; Atr[09,10] = 0 ; Atr[09,13] = 0 ; Atr[10,07] = 0 ;
33  Atr[10,09] = 0 ; Atr[10,13] = 0 ; Atr[13,07] = 0 ; Atr[13,09] = 0 ;
34  Atr[13,10] = 0 ;
35
36  S_Dif_Abs = 1 ;
37
38
39  /*-----*/
40  /* Atribuição para as Caselas não Modeladas (fora de S) com o */
41  /* algoritmo EM, de modo que estas não contribuam com desvios */
42  /* para o modelo de quase-independência ao aplicar a AC sobre a */
43  /* tabela modificada e também possamos obter as ponderações */
44  /* adequadas para linhas e colunas. */
45  /*-----*/
46
47  * Iterações até as Atribuições das Caselas, fora de S, Convergiem ;

```

```

48 DO k = 1 to 100 UNTIL ( S_Dif_Abs < 0.0001 ) ;
49
50 * Total da Matriz de Atribuições ;
51 tAtr = ui' * Atr * ui ;
52
53 * Marginal linha da Matriz de Atribuições ;
54 ma = Atr * ui ;
55
56 * Marginal coluna da Matriz de Atribuições ;
57 mb = Atr' * ui ;
58
59 * Valores a serem Atribuídos para as caselas fora de S ;
60 V0101 = ma[01] * mb[01] / tAtr ; V0202 = ma[02] * mb[02] / tAtr ;
61 V0303 = ma[03] * mb[03] / tAtr ; V0404 = ma[04] * mb[04] / tAtr ;
62 V0505 = ma[05] * mb[05] / tAtr ; V0606 = ma[06] * mb[06] / tAtr ;
63 V0707 = ma[07] * mb[07] / tAtr ; V0808 = ma[08] * mb[08] / tAtr ;
64 V0909 = ma[09] * mb[09] / tAtr ; V1010 = ma[10] * mb[10] / tAtr ;
65 V1111 = ma[11] * mb[11] / tAtr ; V1212 = ma[12] * mb[12] / tAtr ;
66 V1313 = ma[13] * mb[13] / tAtr ; V0709 = ma[07] * mb[09] / tAtr ;
67 V0710 = ma[07] * mb[10] / tAtr ; V0713 = ma[07] * mb[13] / tAtr ;
68 V0907 = ma[09] * mb[07] / tAtr ; V0910 = ma[09] * mb[10] / tAtr ;
69 V0913 = ma[09] * mb[13] / tAtr ; V1007 = ma[10] * mb[07] / tAtr ;
70 V1009 = ma[10] * mb[09] / tAtr ; V1013 = ma[10] * mb[13] / tAtr ;
71 V1307 = ma[13] * mb[07] / tAtr ; V1309 = ma[13] * mb[09] / tAtr ;
72 V1310 = ma[13] * mb[10] / tAtr ;
73
74 * Soma da Diferença absoluta entre os Valores e as anteriores atribuições ;
75 S_Dif_Abs = abs(Atr[01,01] - V0101) + abs(Atr[02,02] - V0202)
76 + abs(Atr[03,03] - V0303) + abs(Atr[04,04] - V0404)
77 + abs(Atr[05,05] - V0505) + abs(Atr[06,06] - V0606)
78 + abs(Atr[07,07] - V0707) + abs(Atr[08,08] - V0808)
79 + abs(Atr[09,09] - V0909) + abs(Atr[10,10] - V1010)
80 + abs(Atr[11,11] - V1111) + abs(Atr[12,12] - V1212)
81 + abs(Atr[13,13] - V1313) + abs(Atr[07,09] - V0709)
82 + abs(Atr[07,10] - V0710) + abs(Atr[07,13] - V0713)
83 + abs(Atr[09,07] - V0907) + abs(Atr[09,10] - V0910)
84 + abs(Atr[09,13] - V0913) + abs(Atr[10,07] - V1007)
85 + abs(Atr[10,09] - V1009) + abs(Atr[10,13] - V1013)
86 + abs(Atr[13,07] - V1307) + abs(Atr[13,09] - V1309)
87 + abs(Atr[13,10] - V1310) ;
88
89 * Novas Atribuições ;
90 Atr[01,01] = V0101 ; Atr[02,02] = V0202 ; Atr[03,03] = V0303 ;
91 Atr[04,04] = V0404 ; Atr[05,05] = V0505 ; Atr[06,06] = V0606 ;
92 Atr[07,07] = V0707 ; Atr[08,08] = V0808 ; Atr[09,09] = V0909 ;
93 Atr[10,10] = V1010 ; Atr[11,11] = V1111 ; Atr[12,12] = V1212 ;
94 Atr[13,13] = V1313 ; Atr[07,09] = V0709 ; Atr[07,10] = V0710 ;
95 Atr[07,13] = V0713 ; Atr[09,07] = V0907 ; Atr[09,10] = V0910 ;
96 Atr[09,13] = V0913 ; Atr[10,07] = V1007 ; Atr[10,09] = V1009 ;
97 Atr[10,13] = V1013 ; Atr[13,07] = V1307 ; Atr[13,09] = V1309 ;
98 Atr[13,10] = V1310 ;
99
100 END ;
101 print 'número de iterações:' k

```

```

102      ', soma das diferenças absolutas:' S_Dif_Abs ;
           K
           S_DIF_AB
número de iterações:      61 , soma das diferenças absolutas: 0.0000921
103
104 * Fim das iterações ;
105
106
107 * Matriz de Atribuições Finais ;
108 print Atr ;
      ATR
      0.1407941      3      3      3      0      2      0      0      1      0
:      1      1      0
      5 394.94675      129      8      3      16      2      23      82      42
:      7      5      32
      2      121 29.021592      21      2      3      0      4      1      3
:      2      11      4
      1      19      18 3.7230494      9      0      0      1      2      1
:      2      15      2
      0      2      4      7 0.2077866      1      0      0      0      0
:      1      1      0
      2      25      2      6      0 1.5240221      0      0      6      2
:      0      2      0
      0      2      0      0      0      0 0.0039731      0 0.1887235 0.1033013
:      0      0 0.0794625
      0      18      2      0      1      1      0 0.8046162      3      3
:      1      2      2
      3      69      0      3      1      16 0.1887235      1 8.9643665 4.9066111
:      2      0 3.7744701
      1      39      1      5      0      2 0.1052878      1 5.0011729 2.7374841
:      3      1 2.105757
      0      7      0      5      0      3      0      2      0      1
: 0.2612187      1      0
      0      16      12      8      2      1      0      0      0      0
:      0 1.1341345      0
      0      33      3      2      0      0 0.077476      1 3.6801083 2.0143751
:      0      0 1.5495193
109
110 * Total da Matriz de Atribuições Finais ;
111 tAtr = ui' * Atr * ui ;
112
113 * Marginal linha da Matriz de Atribuições Finais ;
114 ma = Atr * ui ;
115
116 * Marginal coluna da Matriz de Atribuições Finais ;
117 mb = Atr' * ui ;
118
119
120 * Matriz do modelo de Quase-Independência ;
121 QI = ma * mb' / tAtr ;
122 QI[01,01] = TC[01,01] ; QI[02,02] = TC[02,02] ; QI[03,03] = TC[03,03] ;
123 QI[04,04] = TC[04,04] ; QI[05,05] = TC[05,05] ; QI[06,06] = TC[06,06] ;
124 QI[07,07] = TC[07,07] ; QI[08,08] = TC[08,08] ; QI[09,09] = TC[09,09] ;
125 QI[10,10] = TC[10,10] ; QI[11,11] = TC[11,11] ; QI[12,12] = TC[12,12] ;
126 QI[13,13] = TC[13,13] ; QI[07,09] = TC[07,09] ; QI[07,10] = TC[07,10] ;

```

```

127   QI[07,13] = TC[07,13] ; QI[09,07] = TC[09,07] ; QI[09,10] = TC[09,10] ;
128   QI[09,13] = TC[09,13] ; QI[10,07] = TC[10,07] ; QI[10,09] = TC[10,09] ;
129   QI[10,13] = TC[10,13] ; QI[13,07] = TC[13,07] ; QI[13,09] = TC[13,09] ;
130   QI[13,10] = TC[13,10] ;
131   print QI ;
      QI
      0 7.4569542 2.0214023 0.7141169 0.1812874 0.4632205 0.0236515 0.3365786 1.1234453 0.6142385
: 0.191776 0.399599 0.4730296
  7.4569542 0 107.06066 37.822169 9.6016271 24.533806 1.2526666 17.826402 59.501661 32.56933
: 10.15714 21.164187 25.053331
  2.0214023 107.06066 0 10.252687 2.602772 6.6505294 0.339568 4.8323121 16.12948 8.8287683
: 2.7533583 5.7371059 6.7913602
  0.73403 38.876841 10.538584 0 0.9451423 2.4150008 0.123307 1.7547532 5.8570838 3.2059827
: 0.9998245 2.0833102 2.4661406
  0.1613743 8.5469545 2.3168754 0.818501 0 0.5309305 0.0271087 0.3857771 1.2876619 0.7048255
: 0.2198084 0.4580094 0.5421735
  0.4632205 24.533806 6.6505294 2.3494853 0.596446 0 0.0778147 1.107363 3.6961995 2.0231829
: 0.6309541 1.3147037 1.5562945
  0.0236515 1.2526666 0.339568 0.1199619 0.0304538 0.0778147 0 0.0565406 0 0
: 0.0322158 0.0671272 0
  0.3365786 17.826402 4.8323121 1.7071493 0.433381 1.107363 0.0565406 0 2.6856794 1.4709561
: 0.4584548 0.9552711 1.1308124
  1.1234453 59.501661 16.12948 5.6981896 1.4465562 3.6961995 0 2.6856794 0 0
: 1.5302483 3.1885413 0
  0.6267642 33.195664 8.9985523 3.17899 0.8070261 2.0620902 0 1.4983264 0 0
: 0.8537175 1.7788704 0
  0.191776 10.15714 2.7533583 0.9727007 0.2469322 0.6309541 0.0322158 0.4584548 1.5302483 0.8378096
: 0 0.5442951 0.6443151
  0.399599 21.164187 5.7371059 2.026793 0.5145266 1.3147037 0.0671272 0.9552711 3.1685413 1.7453068
: 0.5442951 0 1.3425437
  0.4612039 24.426998 6.6215762 2.3392568 0.5938494 1.5173872 0 1.1025421 0 0
: 0.6282072 1.3089801 0
132
133
134   * Massas linha a serem consideradas ;
135   a = ma / tAtr ;
136   print a ;
      A
      0.0099566
      0.5273363
      0.1429483
      0.0519087
      0.011412
      0.0327577
      0.0016726
      0.023802
      0.0794471
      0.0443231
      0.0135619
      0.0282586
      0.0326151
137
138   * Massas coluna a serem consideradas ;
139   b = mb / tAtr ;

```

140 print b ;

 B
0.0099566
0.5273363
0.1429483
0.0505005
0.0128202
0.0327577
0.0016726
0.023802
0.0794471
0.0434868
0.0135619
0.0282586
0.0334514

141

142

143 quit ;

Exiting IML.

NOTE: The PROCEDURE IML used 28.00 seconds.

144 run ;

145

146

147

148 /*-----*/

149 /* Fim de Programa */

150 /*-----*/

III.c Listagem do Programa em SAS para ACG, aplicada para a Decomposição dos Resíduos do Modelo de Quase-Independência

```

title 'Etologia em Pássaros';
title2 'Decomposição dos Resíduos do Modelo de Quase-Independência';
title4 'Atribuindo valores às caselas com zeros estrut. (algoritmo EM)';
title5 'Nota: o número de GL válido é 119';

data EPdados;
input
  Ativid$ A B C D E F G H I J L K M @@;
Label
  A = 'A=ESTICA'
  B = 'B=LOCOMO'
  C = 'C=CANTAN'
  D = 'D=ENFEIT'
  E = 'E=COÇAND'
  F = 'F=ESFR.B'
  G = 'G=BANHO '
  H = 'H=ESTERE'
  I = 'I=BEBEND'
  J = 'J=COMEND'
  K = 'K=QUIETO'
  L = 'L=ARREPI'
  M = 'M=AMOLAN';

cards;
a=estica      .1407941      3      3      3      0      2
               0      0      1      0      1      1      0
b=locomo      5 394.94675      129      8      3      16
               2      23      82      42      7      5      32
c=cantan      2      23      121 29.021592      21      2      3
               0      4      1      3      2      11      4
d=enfeit      0      1      19      18 3.7230494      9      0
               0      1      2      1      2      15      2
e=coçand      0      0      2      4      7 0.2077866      1
               0      0      0      0      1      1      0
f=esfr.b      0      2      25      2      6      0 1.5240221
               0      0      6      2      0      2      0
g=banho      0.0039731      0 0.1887235 0.1033013      0      0 0.0794625
               0      0      18      2      0      1      1
h=estere      0 0.8046162      3      3      1      2      2
               3      69      0      3      1      16
i=bebend      0.1887235      1 8.9643665 4.9068111      2      0 3.7744701
               1      39      1      5      0      2
j=comend      0.1052878      1 5.0011729 2.7374841      3      1 2.105757
               0      7      0      5      0      3
k=quieto      0      2      0      1 0.2612187      1      0
               0      0      16      12      8      2      1
l=arrepi      0      0      0      0      0 1.1341345      0
               0      33      3      2      0      0
m=amolán      0.077476      1 3.6801083 2.0143751      0      0 1.5495193;
run;

```

```

title6 'Coordenadas F e G';
proc corresp
  data = EPdados outc = EPout
  dimens = 3
  profile = both
  observed rp cp
  short;
  var A B C D E F G H I J L K M;
  id Ativid;
  run;
proc print data = EPout;
  run;
proc plot
  data = EPout;
  plot dim1*dim2 = Ativid / box href = 0 vref = 0 ;
  plot dim1*dim3 = Ativid / box href = 0 vref = 0 ;
  plot dim2*dim3 = Ativid / box href = 0 vref = 0 ;
  run;

```

```

title6 'Coordenadas  $\Phi$  e  $\Gamma$ ';
proc corresp
  data = EPdados outc = EPout
  dimens = 3
  row = DA column = DB
  short;
  var A B C D E F G H I J L K M;
  id Ativid;
  run;
proc print data = EPout;
  run;

```

```

/*-----*/
/*          Fim de Programa          */
/*-----*/

```

III.d Listagem do Programa em IML/SAS para ACG, aplicada para a Decomposição dos Resíduos do Modelo de Simetria

```
/*-----*/
/* Etologia em Pássaros - ACG para o Modelo de Simetria */
/*-----*/
```

```
proc IML ;
```

```
* Inicialização de Variáveis Auxiliares;
```

```
um = { 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1 };
```

```
* A Tabela de Contingência : N ;
```

```
N = { 0 3 3 3 0 2 0 0 1 0 1 1 0 ,
      5 0 129 8 3 16 2 23 82 42 7 5 32 ,
      2 121 0 21 2 3 0 4 1 3 2 11 4 ,
      1 19 18 0 9 0 0 1 2 1 2 15 2 ,
      0 2 4 7 0 1 0 0 0 0 1 1 0 ,
      2 25 2 6 0 0 0 0 6 2 0 2 0 ,
      0 2 0 0 0 0 0 0 0 0 0 0 0 ,
      0 18 2 0 1 1 0 0 3 3 1 2 2 ,
      3 69 0 3 1 16 0 1 0 0 2 0 0 ,
      1 39 1 5 0 2 0 1 0 0 3 1 0 ,
      0 7 0 5 0 3 0 2 0 1 0 1 0 ,
      0 16 12 8 2 1 0 0 0 0 0 0 0 ,
      0 33 3 2 0 0 0 1 0 0 0 0 0 };
```

```
* Total da Matriz de Contingência : t_N ;
```

```
t_N = um' * N * um ;
```

```
* A Matriz de Assimetria : A ;
```

```
A = ( N - N' ) / ( 2 * t_N ) ;
```

```
print A ;
```

```
* Massas a serem consideradas para as linhas e colunas : vm ;
```

```
vm = ( N + N' ) * um / ( 2 * t_N ) ;
```

```
Print vm ;
```

```
* Matriz a Diagonalizar : Y ;
```

```
Y = diag((vm)##(-1/2)) * A * diag((vm)##(-1/2)) ;
```

```
Z = Y' * Y ;
```

```
print Z ;
```

```
* Cálculo dos Eixos Principais : V ;
```

```
V = eigvec(Z) ;
```

```
Print V ;
```



```

* Cálculo das Inércias Principais : M2 ;
  M2 = eigval(Z) ; M2[13] = 0 ;
  Print M2 ;

* Inércia Total : t_In ;
  t_In = M2' * um ;
  Print t_In ;

* Porções da Inércia contabilizada pelos eixos : p_In_e ;
  p_In_e = M2 / t_In ;
  Print p_In_e ;

* Decomposição da Estatística X2 ao longo dos Eixos : CHI_e ;
  CHI_e = t_N * M2 ;
  Print CHI_e ;

* Estatística X2 Total : t_CHI ;
  t_CHI = t_N * t_In ;
  Print t_CHI ;

* Graus de Liberdade da estatística Chi-Quadrado : GL_CHI ;
  GL_CHI = (um' * um) * (um' * um - 1) / 2 ;
  Print GL_CHI ;

* Cálculo dos Valores Próprios : M ;
  M = (M2)##(1/2) ;
  Print M ;

* Nota : o último eixo não deve ser considerado ;
  M[13] = 1 ; M2[13] = 1 ;
  V[ 1,13] = 0 ;   V[ 2,13] = 0 ;   V[ 3,13] = 0 ;
  V[ 4,13] = 0 ;   V[ 5,13] = 0 ;   V[ 6,13] = 0 ;
  V[ 7,13] = 0 ;   V[ 8,13] = 0 ;   V[ 9,13] = 0 ;
  V[10,13] = 0 ;   V[11,13] = 0 ;   V[12,13] = 0 ;
  V[13,13] = 0 ;

* Cálculo das Coordenadas Principais para as Linhas : F ;
  F = diag((vm)##(-1)) * A * diag((vm)##(-1/2)) * V ;
  Print F ;

* Cálculo das Coordenadas Estandarizadas para as Linhas : FE ;
  FE = F * inv(diag(M)) ;
  Print FE ;

```

```

* Cálculo das Coordenadas Principais para as Colunas : G ;
  G = diag((vm)##(-1/2)) * V * diag(M) ;
  G[7, 1] = 0 ; G[7, 2] = 0 ; G[7, 3] = 0 ; G[7, 4] = 0 ;
  G[7, 5] = 0 ; G[7, 6] = 0 ; G[7, 7] = 0 ; G[7, 8] = 0 ;
  G[7, 9] = 0 ; G[7,10] = 0 ; G[7,11] = 0 ; G[7,12] = 0 ;
  G[7,13] = 0 ;
  Print G ;

* Cálculo das Coordenadas Estandartizadas para as Colunas : GE ;
  GE = G * inv(diag(M)) ;
  Print GE ;

* Tabela de Decomposição da Inércia para o espaço das Linhas : Dec_In_R;
  Dec_In_R = diag(vm) * (F)##2 ;
  print Dec_In_R;

* Contribuições Absolutas Linhas : CA_R;
  CA_R = Dec_In_R * inv(diag(M2)) ;
  Print CA_R;

* Inércia das Linhas : In_R ;
  In_R = Dec_In_R * um ;
  Print In_R;

* Contribuições Relativas às Linhas : CR_R ;
  In_R[7] = 1 ;
  CR_R = inv(diag(In_R)) * Dec_In_R ;
  Print CR_R;
  In_R[7] = 0 ;

* Porções da Inércia contabilizada pelas Linhas : p_In_R ;
  p_In_R = In_R / t_In ;
  Print p_In_R;

* Tabela de Decomposição da Inércia para o espaço das Colunas : Dec_In_C;
  Dec_In_C = diag(vm) * (G)##2 ;
  print Dec_In_C;

* Contribuições Absolutas Colunas : CA_C;
  CA_C = Dec_In_C * inv(diag(M2)) ;
  Print CA_C;

* Inércia das Colunas : In_C ;
  In_C = Dec_In_C * um ;
  Print In_C;

```

```

* Contribuições Relativas às Colunas : CR_C ;
  In_C[7] = 1 ;
  CR_C = inv(diag(In_C)) * Dec_In_C ;
  Print CR_C;
  In_C[7] = 0 ;

* Porções da Inércia contabilizada pelas Colunas : p_In_C ;
  p_In_C = In_C / t_In ;
  Print p_In_C;

```

```

quit ;
run ;

```

```

/*-----*/
/*          Fim de Programa          */
/*-----*/

```