UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP

INSTITUTO DE MATEMÁTICA ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA -IMECC

ÁREA DE ESTATÍSTICA

"MODELOS DE PREFERÊNCIA VIA ESCALONAMENTO MULTIDIMENSIONAL"

Autor : Edwin Moises Marcos Ortega.

Orientadora : Prof. Dra. Regina C. C. P. Moran.

CAMPINAS
Estado de São Paulo - Brasil
Agosto 1997

BIBLIOTECA CENTRAL

Dissertação de Mestrado defendida e aprovada em 20 de agosto de 1997

pela Banca Examinadora composta pelos Profs. Drs.

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

Prof (a). Dr (a). EUGÊNIA TRONCOSO LEONE

Prof (a). Dr (a). CLARICE AZEVEDO DE LUNA FREIRE

"MODELOS DE PREFERÊNCIA VIA ESCALONAMENTO MULTIDIMENSIONAL"

Este exemplar corresponde à redação final da tese, devidamente corrigida e defendida pelo Sr. Edwin Moises Marcos Ortega e aprovada pela comissão Julgadora.

Campinas , 27 de agosto de 1997

Prof. Dra. Regina C. C. P. Moran
Orientadora

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Cientifica, UNICAMP, como requisito parcial para a obtenção do Título de MESTRE em Estatística.

Aos meus pais, Felix e Reyda;
e irmãos, Jaime, Hector e Guisella.

AGRADECIMENTOS

Aos que contribuíram de alguma forma para a realização do presente trabalho;

Agradeço a Prof. Regina C. C. P. Moran, pela orientação, compreensão, colaboração e entusiasmo permanente durante o desenvolvimento do trabalho.

Ao IMECC pela oportunidade concedida e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo suporte financeiro.

Um agradecimento muito especial a José Luis Llanos Carrillo e família, pelo constante apoio e incentivo nas horas mas dificeis e por me fazerem sentir parte de sua família.

Um agradecimento carinhoso a Turma 95 do Mestrado, pelo convívio e intercâmbio nestes anos. Em particular a Elsa Moschetti, Rosemeire Fiaccone, Rui Lyu Komi, Fernando Lucambio, Helymar, Carla Vivacqua e André L. S. de Pinho.

Aos professores, colegas e amigos vinculados de alguma forma à minha formação.

Aos colegas peruanos que, de uma forma ou de outra, incentivaram a conclusão do trabalho, em especial a Javier Ferrua e família e Carlos Huaira.

RESUMO

O trabalho tem por objetivo principal estudar os Modelos de Preferência via Escalonamento Multidimensional, isto é, utilizar as dimensões obtidas mediante a técnica do Escalonamento Multidimensional após o processo de interpretação, as quais são consideradas novas variáveis não observáveis, como variáveis independentes nos Modelos de Preferência. A idéia destes últimos é a partir de modelagem de cada resposta individual de preferência, posicionar o indivíduo no que será chamado espaço dos estímulos. Estas posiciones são conhecidas como Pontos Ideais, onde cada indivíduo, será representado no espaço dos estímulos por um ponto o qual será denominado de Ponto Ideal, indicando assim a preferência máxima do indivíduo pelos estímulos. Neste espaço, pode-se obter agrupamentos de Pontos Ideais em regiões particulares e estes tem papel importante na predição de segmentos no mercado.

Entende-se por Escalonamento Multidimensional técnica multivariada descritiva que permite analisar dados relativos a preferências e semelhanças.

Por outro lado os Modelos de Preferência são processos que permitem analisar, através de experimentos, variáveis cujas respostas expressem preferências individuais.

Outro objetivo do trabalho é divulgar a técnica através de sua aplicação a um conjunto de dados reais de preferência.

ÍNDICE

Pág.

CAPÍTULO I Introdução

1.1.- Introdução Geral 1
1.2.- Esboço de uma Tendência Temporal 7
1.3.- Objetivo e Descrição do Trabalho 9

CAPÍTULO II Escalonamento Multidimensional Métrico (MDS)

2.1.- Introdução 11
2.2.- MDS Métrico a Partir de uma Matriz de Dados 12
2.2.1.- Teorema do MDS Métrico 14
2.2.2.- Consequencias do Teorema do MDS Métrico 21
2.2.3.- Algoritmo Prático 22
2.2.4.- Propriedades Ótimas do MDS 23
2.2.5.- Bondade de Ajuste 26
2.3.- MDS Métrico a Partir de uma Matriz de Similaridades 27
2.3.1.- Análise de uma Matriz de Similaridades 27
2.3.2.- Teorema Sobre uma Matriz de Similaridades 28
2.3.3.- Conseqüências do Teorema Sobre uma Matriz de Similaridades 31
2.3.4.- Bondade do Ajuste 32
2.4.- MDS Métrico de Quadrados Mínimos 33

CAPÍTULO III Outros Tipos de Escalonamento Multidimensional.

3.1.- Introdução 35
3.2.- Escalonamento Multidimensional Não Métrico 36
3.2.1.- Problema Geral do MDS Não Métrico 36
3.2.2.- Solução de MDS não Métrico 37

3.2.3.- Algoritmo de Shepard-Kruskal 43
3.3.- Modelo Euclidiano Ponderado 46
3.3.1.- Problema Geral do Modelo Euclidiano Ponderado 47
3.3.2.- Solução do Modelo Euclidiano Ponderado 48
3.3.3.- Ajustando o Modelo Euclidiano Ponderado 49
3.3.4.- Escolha da Dimensão e Bondade do Ajuste 51
3.4.- MDS Métrico a Partir de uma Matriz Retangular de Dados 53
3.5.- Generalização do Modelo Euclidiano Ponderado 54

CAPÍTULO IV Modelos de Preferência

4.1.- Introdução 57
4.2.- Problema Geral dos Modelos de Preferência 58
4.3.- Definição dos Modelos de Preferência 59
4.3.1.- Modelo de Preferência Vetor 59
4.3.2.- Modelos de Preferência Ponto Ideal ou Euclidiano 61
4.3.2.1.- Modelo Ponto Ideal Simples 61
4.3.2.2.- Modelo Ponto Ideal Ponderado 63
4.3.2.3.- Modelo Ponto Ideal Geral 65
4.4.- Estimação dos Modelos de Preferência 72
4.4.1. Estimação Métrica 72
4.4.1.1.- Estimação do Modelo de Preferência Vetor 72
4.4.1.2.- Estimação dos Modelos Ponto Ideal 73
4.4.2. Estimação Não Métrica 84
4.4.2.1.- Algoritmo da Estimação Não Métrica 85
4.4.2.2.- Critério de Convergência 86
4.5.- Observações a respeito dos Modelos de Preferência 86

CAPÍTULO V Aplicação em Mercadologia

5.1.- Introdução ao Problema 87
5.2.- Coleta dos Dados 88
5.3.- Metodologia Estatística 91

5.4.- Resultados via MDS para Cada Grupo 93
5.5.- Resultados via MDS para o Total sem Considerar Estrutura de Grupos 103
5.6.- Resultados via MDS Considerando Estrutura de Grupos 106
5.7.- Análise dos Resultados Obtidos através do MDS 112
5.8.- Resultados dos Modelos de Preferência 114
5.9.- Análise dos Resultados dos Modelos de Preferência 128
5.10.- Resultados Gerais da Aplicação 129

APÊNDICES

- Apêndice A : Conceitos de Distância e Similaridades 130
- Apêndice B : Método Q e Método R na Análise Multivariada 145
- Apêndice C : Estrutura de Matrizes 148
- Apêndice D : Relação entre Produto Escalar e Distância 157
- Apêndice E : Regressão Monotônica 165
- Apêndice F : Níveis de Mensuração ou Medida 171
- Apêndice G : Programas Utilizados Através do SAS 175

Capítulo I

INTRODUÇÃO

1.1.- Introdução Geral

Na atualidade a Análise Multivariada tem sido aplicada em diferentes áreas de pesquisa, por exemplo, em mercadologia[1], onde tem ocasionado uma revolução. Na mercadologia usualmente são utilizados os métodos clássicos da estatística descritiva (que estudam cada variável em separado) , mas o mercado é um fenômeno complexo sobre qual incide grande número de fatores, que precisam ser estudados simultaneamente, e isto consegue-se através das técnicas multivariadas.

Hoje em dia, graças aos computadores, as empresas podem dispor facilmente de grande quantidade de dados referentes à sua atividade, e uma boa análise destes pode diminuir significativamente os riscos das decisões.

Devido às técnicas multivariadas permitirem analisar de forma simultânea duas ou mais variáveis observadas, elas dão uma visão em conjunto dos fenômenos, desenvolvendo a complexa interação dos fatores que existe nos fenômenos da mercadologia e, neste sentido, as técnicas constituem um avanço frente aos estudos clássicos univariados.

A seguir apresentamos alguns aspectos que são a base da aplicação da Análise Multivarida em mercadologia:

- Permite analisar grande número de questionários, os quais proporcionam às empresas grande volume de informação sobre os mercados atuais ou futuros.

[1] Mercadologia.- Relacionada com a pesquisa e estudo do mercado.

- Capacidade de tornar assimilável e compreensível, para a mente humana, grande quantidade de dados com uma perda mínima de informação, proporcionando, em muitos casos, uma imagem gráfica.
- Permite analisar simultaneamente a informação sobre o fenômeno, considerando um número adequado dos possiveis fatores que nele intervêem.
- Permite utilizar tanto variáveis contínuas como discretas, como mistura destas.

Neste contexto é que surge a idéia de realizar este trabalho, o qual é abordar duas técnicas multivariadas pouco conhecidas, o Escalonamento Multidimensional (MDS) e os Modelos de Preferência, articuladamente. Tudo isto, baseado no trabalho apresentado por Warren (1992) sobre mercadologia. Entende-se por Escalonamento Multidimensional, uma técnica multivariada descritiva que permite analisar dados relativos a preferências e semelhanças; assim, também os Modelos de Preferência são processos que permitem obter e analisar, através da estimação de modelos, experimentos cujas variáveis respostas expressem preferências individuais. Porém as técnicas multivariadas diferenciam-se e desenvolvem-se para atender objetivos específicos dependendo da natureza dos dados. Assim, por exemplo, a técnica de componentes principais trata de reduzir a dimensão em variáveis contínuas a partir de uma matriz de dados. A análise discriminante trata de obter combinações lineares através das variáveis e tem por objetivo o estudo do padrão de diferenciação de distintos grupos, com o objetivo de classificação. Assim por diante poderíamos citar cada técnica com seu objetivo e estrutura de dados. Neste contexto rico em metodologia de análise multivariada cabe entender a especificidade tanto do Escalonamento Multidimensional como dos Modelos de Preferência. A diferença entre estas duas últimas técnicas multivariadas e as técnicas clássicas é em relação a natureza dos dados e as suas aplicações. Com este objetivo, introduzimos a seguir alguns termos pouco usuais na análise multivariada:

- Objeto: Um objeto qualquer. Por exemplo, uma laranja.
- Estímulo: Um objeto percebido. Por exemplo, uma laranja saboreada.
- Atributo: Relacionado com a característica percebida de um estímulo, por exemplo: a laranja está doce.
- Proximidade: Um número que mostra o grau de similaridade ou diferença entre um par de estímulos.

Com respeito à natureza dos dados, tanto o MDS quanto os Modelos de Preferência têm formas específicas de coletar os dados em relação às técnicas multivariadas clássicas. Em relação à escolha dos estímulos, não existe uma regra geral, pois dependendo do estudo isso pode variar. No entanto, deve-se ter cuidado no momento de selecionar os estímulos que realmente sejam importantes, pois um número alto de estímulos pode dificultar a avaliação do

entrevistado, prejudicando a credibilidade dos resultados. Na literatura existem vários métodos de julgamentos quantitativos. Baseados no trabalho de Green e Tull (1978) mencionaremos alguns deles. Os mais comuns e usados são: método de julgamentos diretos, método de fracionamentos, método de soma constante e ordenamento através de postos. Como visto em toda análise estatística há que considerar a metodologia de coleta dos dados, e neste caso com cuidados adicionais. A seguir, na descrição detalhada dos quatro métodos, é interesante notar que não ha medidas explícitas das variáveis em estudo mais sim comparação inter-estímulos ou intra-estímulos. Em cada caso na resposta do indivíduo, distintos graus de habilidade aritmética são necessários para refletir, reconhecidamente, o complexo processo de julgar até as próprias preferências ou atitudes.

1- Métodos de julgamentos diretos

Nos métodos de julgamentos diretos, os estímulos são avaliados um a um a partir de uma escala de preferência ou semelhança. Este método pode ser dividido em dois casos:

1.1- Categoria de resposta ilimitada (escala contínua): Neste caso, a escala, pode ser obtida ligando as duas frases com um traço e pedindo para que o entrevistado assinale nessa linha, a posição que mais concorda com sua preferência ou semelhança, seguindo o critério de proximidade das frases (quanto mais próximo de uma frase, maior a concordância com ela). A avaliação é dada através da distância entre uma das frases e o sinal (ou marca) feito pelo entrevistado.

Nenhuma preferência |----------*-----| Altamente preferido

ou analogamente:

Nenhuma semelhança |------*----------| Altamente semelhante

Como se pode observar a decisão do entrevistado é uma resposta livre. Este tipo de avaliação, às vezes, é considerado um tipo de avaliação gráfica.

1.2.- Categoria de resposta limitada (escala discreta): Aqui, as respostas são limitadas para escolher a preferência ou semelhança. Por exemplo, ao utilizar uma escala de 7 pontos, são apresentados aos entrevistados números variando de 1 a 7. Ao número 7 pode-se associar a frase "altamente preferido", assim pede-se ao entrevistado que atribua, para cada estímulo, um

número de 1 a 7 de tal modo que quanto maior o número, maior é a preferência pelo estímulo, isto é:

Pouca preferência.	1	2	3	4	5	6	7	Maior preferência.

2.- Método de fracionamento

O método de fracionamento é um tipo de procedimento de julgamentos quantitativos, no qual as respostas são dadas em relação a dois estímulos. Pode ser feito de três formas diferentes:

2.1.- Estímulo principal: Usado quando existe um estímulo principal (ou atraente) em estudo. Por exemplo, se temos 5 estímulos (A,B,C,D,E) e, estamos interessados em analisar o estímulo "D" contra os demais, então pede-se ao entrevistado para comparar cada um dos estímulos A, B, C e E com o D, com respeito a algum atributo, onde o atributo do estímulo D é assumido com uma pontuação de 1.00. Uma resposta hipotética poderia ser:

ESTÍMULO	RESPOSTA
A	0.75
B	0.80
C	0.40
E	0.50

2.2.- Comparação pareada: Neste caso, os estímulos são apresentados aos pares, e o entrevistado deve escolher o preferido. Um método alternativo de avaliação seria pedir que o entrevistado além de escolher o estímulo preferido, indicasse também, a intensidade da diferença ou semelhança entre os dois estímulos.

2.3- Mistura: Seria misturar o item (2.2) com o item (1), isto é, para cada par de estímulos dar uma preferência ou semelhança. Por exemplo se temos 5 estímulos (A, B, C, D, e E), podemos formar 10 pares de estímulos. Logo pede-se ao entrevistado dar uma escala de preferência ou semelhança entre cada par de estímulos, para fornecer esta escala podem ser utilizados a categoria de resposta ilimitada (ver 1.1) ou categoria de resposta limitada (ver 1.2).

3.- Método de Soma Constante

O Método de soma constante, por sua simplicidade de instruções, é muito usado em mercadologia. Neste método, o entrevistado é perguntado com respeito a todos os estímulos e tem que dar uma nota em uma escala de 100 (ou 10 pontos) de acordo como a sua preferência. Por exemplo, se temos 5 estímulos (A,B,C,D,E) e se perguntamos ao entrevistado: Distribua 100 pontos através dos 5 estímulos de acordo com o grau relativo de sua preferência para os respectivos estímulos. Uma resposta poderia ser a seguinte:

ESTÍMULO	RESPOSTA
A	20
B	25
C	10
D	5
E	40
total	100

4.- Ordenamento através de postos

Outra forma de trabalhar é através da ordenação dos postos. Esta avaliação consiste na ordenação dos estímulos segundo a preferência ou semelhança do entrevistado. Recomenda-se que a ordenação seja em ordem crescente. Por exemplo, suponha que haja 5 estímulos A, B, C ,D e E, e eles sejam apresentados ao entrevistado. Em seguida, pede-se ao entrevistado que dê uma ordem de preferência para esses estímulos. Suponha que o entrevistado dê a seguinte sequência de preferência D, B, C, E e A (da maior para a menor preferência). Assim, para os estímulos A, B, C, D e E haverá a seguinte atribuição de postos 5, 2, 3, 1, 4, respectivamente, se a atribuição for em ordem decrescente, e 1, 4, 3, 5, 2 se for em ordem crescente de preferência.

As técnicas do MDS e os Modelos de Preferência podem ser aplicados para analisar matrizes de distâncias, similaridades ou correlações. Dillon e Goldstein (1984) argumentam que o MDS é muito usado em diferentes áreas de pesquisas. A seguir são apresentadas uma série de aplicações ilustrativas (com maior ênfase na mercadologia):

a.- Cientistas políticos têm usado a técnica de MDS frequentemente para estudar o comportamento dos candidatos políticos, assim como para analisar o ranking de popularidade.

b.- Antropólogos têm usado muito a técnica de MDS em estudos de diferentes culturas de vários grupos humanos, baseados em opinião, linguagem e outras informações históricas.

c.- Projetos urbanos e regionais têm procurado identificar similaridades entre várias cidades, prefeituras, municípios ou regiões a fim de estudar a posição e obter uma configuração de pontos em um espaço reduzido através do MDS. Iniciando a análise a partir de dados demográficos, fiscais ou econômicos.

d.- Psicólogos têm usado esta classe de procedimentos para entender e perceber a avaliação das características de conversa (modo de falar), cores e personalidade, entre outras coisas.

e.- Pesquisadores da área da Análise Sensorial têm usado muito o MDS para analisar o comportamento de diferentes atributos de um produto.

Nesta parte da aplicação, estamos mais interessados no que acontece em relação à mercadologia, pois a aplicação do capítulo V será feita nessa área de pesquisa. A seguir abordaremos alguns campos específicos de aplicação do MDS, baseados no trabalho apresentado por Cambillo (1996).

As aplicações fundamentais do MDS, na área de mercadologia, servem, principalmente, para analisar:

i.- As percepções do consumidor sobre a semelhança e a preferência de marcas (ou estímulos).
ii.- As preferências do consumidor pelas marcas (ou estímulos).

Neste sentido, o objetivo da mercadologia é a obtenção de informações para poder adotar decisões comerciais corretas. Pode desmembrar-se em outros vários objetivos por áreas ou atividades, que dão lugar a múltiplos objetivos específicos, sendo que alguns deles são:

a) Investigação sobre o comportamento do consumidor: Motivações de compra, pessoas que influem e decidem na compra, atitudes e intenções de consumidores, hábitos de compra, estilos de vida e segmentação dos consumidores.

b) Investigação sobre a demanda e vendas: Determinação da demanda total do mercado de um produto, determinação da demanda total de um grupo de produtos afins ou de um setor, participação das marcas no mercado, determinação da demanda por zonas geográficas e tipos de consumidores, determinação de índices de capacidade de compra, previsões de vendas.

c) Investigação sobre os produtos: Criação de novos produtos, modificação de produtos, eliminação de produtos, teste de conceito de novo produto, teste de produto, teste de mercado, teste de nome, imagem e posicionamento de marcas.

d) Investigação sobre a publicidade: Influência da publicidade no comportamento dos consumidores, determinação da audiência dos meios de comunicação, medição da eficácia publicitária, relação entre investimento publicitário e vendas.

e) Investigação sobre a promoção de vendas: Atitudes do consumidor final, avaliação ou escolha dos programas promocionais, comparação entre os resultados da publicidade e promoção, controle dos resultados da promoção.

f) Investigação sobre a distribuição: Tarefas necessárias para pôr nos pontos de venda os produtos fabricados pela empresa, logística dos canais de distribuição e a organização da distribuição física.

1.2.- Esboço de uma Tendência Temporal.

A origem do Escalonamento Multidimensional, vista como a identificação de seus precursores é um tanto confusa. A formulação da metodologia acontece independentemente por vários autores, em contextos diferentes. No entanto, existe um consenso em considerar a Young e Householder (1938) como os iniciadores, com o estudo baseado em estrutura de matrizes, aproximação de matrizes e relação entre distância e produto escalar. Paralelamente, Richardson (1938) propõe analogia entre conceito psicológico de similaridade e conceito geométrico de distância, a qual serviu de base para que Torgerson (1952, 1958) iniciasse suas pesquisas, introduzindo pela primeira vez o termo MDS Métrico, e utilizando basicamente a distância Euclidiana. Shepard e Kruskal (1962, 1964) introduzem o termo de MDS não métrico, onde a análise é baseada nos postos das distâncias ou similaridades, e usam métodos iterativos para solucionar o problema. No entanto, Gower (1966) baseado na idéia de

Torgerson, analisa a técnica a partir de uma matriz de similaridades, a qual é conhecida com o nome de Coordenadas Principais. Também, Coombs (1964) propôs modelos de distância para dados de preferência conhecidos como Modelos Desdobrados (Unfolding Model), os quais servirão de base para os Modelos de Preferência. Horan (1969), Caroll e Chang (1970) descrevem pela primeira vez o MDS sob estrutura de grupos, ao qual denominarão de Escalonamento de Diferenças Individuais (Individual Differences Scaling), também conhecido com o nome de Modelo Euclidiano Ponderado (The Weighted Euclidean Model). Lester (1969) apresenta um trabalho sobre aplicação do Escalonamento Multidimensional não Métrico na mercadologia. Caroll (1972) e Tucker (1972) estudam a Generalização do Modelo Euclidiano Ponderado.

Davison (1983) menciona que Caroll (1972) divide a análise de dados de preferência em duas partes:

- MDS Interno : Onde, basicamente, o objetivo é estimar uma configuração de pontos a partir de uma matriz de distâncias ou similaridades. O estudo é dirigido para analisar dados de preferência ou semelhança.

- MDS Externo : Baseia-se em formular Modelos de Preferência através de distâncias. O estudo está dirigido, especialmente, para dados de preferências individuais.

A partir da década de 80, a maioria dos trabalhos deram ênfase no melhoramento da parte computacional do MDS. Um importante trabalho é apresentado por Takane (1987), que desenvolveu um método alternativo para analisar tabelas de contingência chamado de Análise Discriminante Ponto Ideal. Este consiste da combinação de 4 técnicas conhecidas: MDS, Modelo Loglinear, Análise de Correspondência e Análise Canônica. Por outro lado, Caroll (1989) apresenta um método quase métrico para MDS via uma extensão do Modelo Euclidiano.

Já nos anos 90, Warren (1992) apresenta um trabalho sobre as técnicas multivariadas em mercadologia. Por outro lado, Rousseeuw e Molenberghs (1993) utilizam a idéia de MDS não métrico para transformar matrizes de correlação semidefinida não positiva[2]. Uma importante aplicação conjunta do MDS métrico e MDS não métrico é apresentada por Falahee

[2] Matriz semi definida não positiva.- Quando as formas quadráticas associadas à dita matriz são maior ou igual a zero.

e MacRae (1995) para avaliar o consumo de 13 diferentes tipos de água. Em 1995 são realizados dois trabalhos importantes. Groenen (1995) estuda uma aproximação melhorada para o MDS usando a distância de Minkowski e Carroll, Mathar e Heiser (1995) ajusta e estende o modelo de Escalonamento de Diferenças Individuais.

1.3.- Objetivo e Descrição do trabalho.

O presente trabalho tem por objetivo principal estudar os Modelos de preferência via Escalonamento Multidimensional, isto é, utilizar as dimensões obtidas mediante a técnica do MDS após o proceso de interpretação, as quais são consideradas novas variáveis não observáveis, fatores ou atributos, etc, como variáveis independentes nos Modelos de Preferência. A idéia destes últimos é, a partir da modelagem de cada resposta individual de preferência, posicionar o indivíduo no que será chamado espaço dos estímulos[3]. Estas posições são conhecidas como Pontos Ideais, onde cada indivíduo será representado no espaço dos estímulos por um ponto, o qual será denominado de ponto ideal, indicando assim a preferência máxima do indivíduo pelos estímulos. Neste espaço, pode-se obter agrupamentos de pontos idéais em regiões particulares e estes tem papel na predição de segmentos no mercado.

O objetivo principal do trabalho foi baseado na publicação feito por de Warren (1992), no qual encontramos alguns métodos estatísticos para analisar dados referentes a mercadologia, assim como a implementação destes métodos no softaware estatístico SAS. Os métodos mais comuns e apropriados que sugere Warren são: Modelos de Preferência (Conjoint Analysis), Análise de Correspondência (Correspondence Analysis), Mapas de Preferência (Preference Mapping), Análise de Preferência Multidimensional (Multidimensional Preference Analysis) e Escalonamento Multidimensional (Multidimensional Scaling). Por outro lado Warren sugere utilizar a solução do Escalonamento Multidimensional para formular os Modelos De Preferência. Logicamente para realizar este objetivo o primeiro passo foi estudar e analisar as técnicas multivariadas do Escalonamento Multidimensional assim como os Modelos de Preferência.

Assim, o presente trabalho propõe-se a ser um roteiro que permita a um pesquisador aplicar o MDS articuladamente aos Modelos de Preferência. Estão descritas as várias etapas que devem ser percorridas para se aplicar os Modelos de Preferência via MDS e como

[3] Espaço dos estímulos.- Esta denominação deriva do fato que os eixos são as novas variáveis não observáveis da solução do MDS e, portanto, propiciam o espaço dos estímulos.

classificar os diversos tipos de MDS, para finalmente fazer uma análise de preferência, abrangendo desde a coleta dos dados até uma análise dos resultados.

Para alcançar nosso objetivo, descrevemos, a seguir, a estrutura do presente trabalho. No capítulo II há uma breve análise algébrica do MDS Métrico, dando maior destaque ao método clássico de MDS através de uma matriz de dados e de uma matriz de similaridades. Finalmente apresentamos, em forma resumida, o MDS métrico de quadrados mínimos.

O capítulo III abrange uma descrição do MDS não métrico, o qual está baseado na idéia de Shepard e Kruskal (1962, 1964). Mencionamos também o MDS sob estrutura de grupos, conhecido com o nome de Escalonamento de Diferenças Individuais (Individual Differences Scaling) ou Modelo Euclidiano Ponderado (The Weighted Euclidean Model), assim como também a forma de analisar uma matriz retangular de preferências individuais, para, ao final desse capítulo, tratar numa forma simplificada a Generalização do Modelo Euclidiano Ponderado.

No capítulo IV, há um relato da Análise de Preferência, através de seus quatro modelos mais comuns: Modelo de Preferência Vetor e os Modelos Pontos Ideais Simples, Ponderado e Geral. Também descrevemos a forma métrica e não métrica de estimar os modelos.

No capítulo V, é feita a aplicação na área de mercadologia, utilizando dados reais de preferências individuais, cedidos por uma empresa de pesquisa de mercado, os quais fizeram parte de um projeto de uma entidade financeira que, por razões de confidencialidade, não terá o nome mencionado.

Capítulo II

ESCALONAMENTO MULTIDIMENSIONAL MÉTRICO

2.1.- Introdução

A técnica de Escalonamento Multidimensional Métrico (MDS métrico), enfoca o problema de construir uma configuração de n pontos em um espaço euclidiano, usando como informação principal distâncias ou similaridades entre n objetos ou indivíduos. Estas distâncias ou similaridades são representadas em uma matriz, denominada matriz de distâncias ou matriz de similaridades, e a solução do MDS Métrico envolve a autoestrutura de tal matriz.

Existem dois aspectos que caracterizam o MDS Métrico. O primeiro deles é o nível de mensuração dos dados (ver apêndice F). Se eles estiverem em escala razão, portanto satisfazendo as propriedades de distâncias ou na escala intervalar, podendo ser transformadas em distâncias mediante a adição de uma constante, o MDS é métrico. O segundo aspecto diz respeito ao procedimento utilizado para a determinação da configuração. O MDS métrico admite a existência de uma verdadeira configuração de pontos em k dimensões, cujas distâncias entre os pontos são d_{rs}. Pretende-se reconstruir esta configuração usando a matriz de distâncias observadas $\Delta^2 = (\delta_{rs})$ cujos elementos são da forma:

$$\delta_{rs} = d_{rs} + e_{rs}$$

onde e_{rs} representam erros de medida, acrescidos de erros de distorções, pelo fato das distâncias observadas não corresponderem exatamente às distâncias de uma configuração em R^k.

Neste capítulo, será feita a apresentação formal do Escalonamento Multidimensional Métrico (MDS métrico), trabalhando basicamente com distância Euclidiana e usando uma

matriz de projeção. Esta apresentação, será realizada através de dois casos: o primeiro caso é quando temos uma matriz de dados (X), e o segundo caso quando temos uma matriz de similaridades (S). No primeiro caso, a partir de uma matriz de dados, calcularemos distâncias (entre indivíduos ou variáveis), para estudar essas distâncias, isto é, encontrar uma nova configuração de pontos em menor dimensão, de tal forma que reproduzam as distâncias originais. Por esta razão, o MDS métrico é considerado como uma técnica multivariada que reduz dimensão. Também serão apresentadas propriedades ótimas, assim como medida de bondade do ajuste.

2.2.- MDS Métrico a partir de uma Matriz de Dados

A análise do MDS métrico foi desenvolvida teoricamente nos fundamentos do Escalonamento Clássico (Scaling Classico) proposto por Young e Householder (1938) (ver apêndices C e D), que é um método algébrico para relacionar produto escalar e distâncias.

Nos estudos das técnicas da Análise Multivariada é comum ter uma matriz de dados X_{nxp} formada por n observações, indivíduos ou casos medidos (mensurados) em p variáveis (p<n). Tal matriz é definida da seguinte maneira:

$$X_{nxp} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{matrix} \vec{X}_1^t \\ \vec{X}_2^t \\ \vdots \\ \vec{X}_n^t \end{matrix} \qquad (2.2.1)$$

A partir da matriz de dados X podemos obter distâncias entre indivíduos ou distâncias entre variáveis. Nesta seção trabalhamos basicamente com os indivíduos. A distância Euclidiana entre os vetores $\vec{X}_i$ e $\vec{X}_j$ (isto é, entre os indivíduos i e j) é dado por (ver apêndice A):

$$d_{ij}^2 = \sum_{k=1}^{p} \left(x_{ik} - x_{jk}\right)^2 = \left\|\vec{X}_i - \vec{X}_j\right\|^2 \qquad (2.2.2)$$

Como temos n indivíduos então, teremos $1/2(n)(n-1)$ distâncias, lembrando que a distância para i=j é zero. Estas distâncias podem ser dispostas em uma matriz da seguinte maneira:

$$\Delta^2_{nxn} = \begin{pmatrix} 0 & d^2_{12} & \cdots & d^2_{1n} \\ d^2_{21} & 0 & \cdots & d^2_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d^2_{n1} & d^2_{n2} & \cdots & 0 \end{pmatrix}_{nxn} \tag{2.2.3}$$

DEFINIÇÃO 2.2.1

Uma matriz Δ^2 de ordem (nxn) mostrada na equação (2.2.3) é chamada de MATRIZ DE DISTÂNCIAS, se é simétrica, $d_{rr} = 0$ e $d_{rs} \geq 0$ para todo r,s=1,2,...,n. e r≠s.

DEFINIÇÃO 2.2.2

Uma matriz de distâncias é Euclidiana se existe alguma configuração de pontos $\vec{X}_1, \vec{X}_2, .., \vec{X}_n \in R^p$, tal que a distância entre 2 pontos quaisquer, $\vec{X}_r$ e $\vec{X}_s$, é:

$$d^2_{rs} = \sum_{k=1}^{p} (x_{rk} - x_{sk})^2 = \|\vec{X}_r - \vec{X}_s\|^2 = (\vec{X}_r - \vec{X}_s)^t (\vec{X}_r - \vec{X}_s) \tag{2.2.4}$$

Então, usando estas definições pode-se observar que a matriz (2.2.3) é uma matriz de distâncias Euclidianas.

Agora que temos uma matriz com distâncias Euclidianas, estudaremos sua auto-estrutura. Para isto utilizaremos os apêndices B e C e a explicação teórica de Mardia, Kent e Bibby (1979).

Para alguma matriz de distâncias Δ^2 (não necessariamente Euclidiana) seja:

$$A=[a_{rs}] \quad \text{onde:} \quad a_{rs}=(-1/2)d_{rs}^2 \tag{2.2.5}$$

que na forma matricial é dado por:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}_{nxn} = \begin{pmatrix} 0 & -\frac{d_{12}^2}{2} & \cdots & -\frac{d_{1n}^2}{2} \\ -\frac{d_{21}^2}{2} & 0 & \cdots & -\frac{d_{2n}^2}{2} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{d_{n1}^2}{2} & -\frac{d_{n2}^2}{2} & \cdots & 0 \end{pmatrix}_{nxn}$$

Pela transformação de Dupla Centralização (ver apêndice D), temos o seguinte:

$$B_{nxn}=H_{nxn}\,A_{nxn}\,H_{nxn} \tag{2.2.6}$$

onde: $\mathbf{H}=\mathbf{I}-(1/n)\mathbf{1}\mathbf{1}^t$ (matriz que projeta ortogonalmente o espaço sobre o subespaço ortogonal a $\mathbf{1}$) e $\mathbf{1}=[1,1,...,1]^t$.

Podemos notar que a matriz $\mathbf{H}$ de projeção é um caso particular da matriz θ (definida no apêndice C) .

O próximo teorema é conhecido como Teorema Fundamental de MDS assinalado assim por Davison (1983). Este teorema garante que a partir de uma matriz de distância Euclidiana é possível construir uma matriz B semi definida positiva, e da matriz B é possível construir uma configuração de pontos através dos auto-vetores associados aos auto-valores não nulos dessa matriz de tal forma que, se calcularmos as distâncias Euclidianas dessa configuração de pontos, estas reproduzirão exatamente as distâncias Euclidianas originais. O teorema foi desenvolvido por Torgerson (1952) baseado na idéia de Young-Householder (1938), e depois foi analisado por Mardia, Kent e Bibby (1979), Davison (1983), Krzanowski (1995) e outros.

2.2.1.-Teorema do MDS Métrico

Teorema 2.2.1.1

Dada uma matriz de distâncias Δ^2, definimos B como a transformação de dupla centralização (ver equação 2.2.6), então a matriz Δ^2 é Euclidiana, se e somente se, B é uma matriz semi-definida positiva (s.d.p).

Prova(⇒)

Hipótese: A matriz Δ^2 é Euclidiana. Isto implica que existe uma configuração de pontos $\vec{X}_1,\vec{X}_2,...,\vec{X}_n \in R^p$ tal que:

$$\begin{aligned} d_{rs}^2 &= \sum_{k=1}^{p}\left(x_{rk}-x_{sk}\right)^2 \\ &= \left\|\vec{X}_r-\vec{X}_s\right\|^2 \\ &= \left(\vec{X}_r-\vec{X}_s\right)^t\left(\vec{X}_r-\vec{X}_s\right) \\ &= \vec{X}_r^t\vec{X}_r-2\vec{X}_r^t\vec{X}_s+\vec{X}_s^t\vec{X}_s \\ &= -2a_{rs} \end{aligned}$$

Por outro lado, sabe-se que:

$$B_{nxn}=H_{nxn}\,A_{nxn}\,H_{nxn}=[\,I-(1/n)\mathbf{1}\mathbf{1}^t]\,A\,[\,I-(1/n)\mathbf{1}\mathbf{1}^t]$$

Então, temos que:

$$B = A-(1/n)\mathbf{1}\mathbf{1}^tA-(1/n)\,A\,\mathbf{1}\mathbf{1}^t+(1/n^2)\mathbf{1}\mathbf{1}^tA\,\mathbf{1}\mathbf{1}^t$$

Sejam:

i) $(1/n)\mathbf{1}\mathbf{1}^tA$

ii) $(1/n)\,A\,\mathbf{1}\mathbf{1}^t$

iii) $(1/n^2)\mathbf{1}\mathbf{1}^tA\,\mathbf{1}\mathbf{1}^t$

Expressando i, ii e iii na forma matricial temos:

$$i)\rightarrow n^{-1}\mathbf{1}\mathbf{1}^tA=\begin{pmatrix} \bar{a}_{.1} & \bar{a}_{.2} & \cdots & \bar{a}_{.n} \\ \bar{a}_{.1} & \bar{a}_{.2} & \cdots & \bar{a}_{.n} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{.1} & \bar{a}_{.2} & \cdots & \bar{a}_{.n} \end{pmatrix}_{nxn} \quad \Rightarrow \quad \bar{a}_{.r}=\frac{1}{n}\sum_{s=1}^{n}a_{rs}$$

$$ii) \rightarrow n^{-1}\mathbf{A}\mathbf{1}\mathbf{1}^t = \begin{pmatrix} \bar{a}_{1.} & \bar{a}_{1.} & \cdots & \bar{a}_{1.} \\ \bar{a}_{2.} & \bar{a}_{2.} & \cdots & \bar{a}_{2.} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{n.} & \bar{a}_{n.} & \cdots & \bar{a}_{n.} \end{pmatrix}_{nxn} \quad \Rightarrow \quad \bar{a}_{s.} = \frac{1}{n}\sum_{r=1}^{n} a_{rs}$$

$$iii) \rightarrow n^{-2}\mathbf{1}\mathbf{1}^t\mathbf{A}\mathbf{1}\mathbf{1}^t = \begin{pmatrix} \bar{a}_{..} & \bar{a}_{..} & \cdots & \bar{a}_{..} \\ \bar{a}_{..} & \bar{a}_{..} & \cdots & \bar{a}_{..} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{..} & \bar{a}_{..} & \cdots & \bar{a}_{..} \end{pmatrix}_{nxn} \quad \Rightarrow \quad \bar{a}_{..} = \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}$$

então cada elemento da matriz **B** pode ser escrito da seguinte maneira:

$$b_{rs} = a_{rs} - \bar{a}_{.r} - \bar{a}_{s.} + \bar{a}_{..} \tag{2.2.7}$$

Por outro lado, da expressão i, temos:

$$\begin{aligned}
\bar{a}_{.r} &= \frac{1}{n}\sum_{s=1}^{n} a_{rs} \\
&= \frac{1}{n}\left[a_{r1} + a_{r2} + \ldots + a_{rn}\right] \\
&= \frac{1}{n}\left[\left(-\tfrac{1}{2}\right)d_{r1}^2 + \left(-\tfrac{1}{2}\right)d_{r2}^2 + \ldots + \left(-\tfrac{1}{2}\right)d_{rn}^2\right] \\
&= -\frac{1}{2n}\left[d_{r1}^2 + d_{r2}^2 + \ldots + d_{rn}^2\right] \\
&= -\frac{1}{2n}\begin{bmatrix}\left(\vec{X}_r - \vec{X}_1\right)^t\left(\vec{X}_r - \vec{X}_1\right) + \left(\vec{X}_r - \vec{X}_2\right)^t\left(\vec{X}_r - \vec{X}_2\right) \\ + \ldots + \left(\vec{X}_r - \vec{X}_n\right)^t\left(\vec{X}_r - \vec{X}_n\right)\end{bmatrix} \\
&= -\frac{1}{2n}\begin{bmatrix}\left(\vec{X}_r^t\vec{X}_r + \vec{X}_1^t\vec{X}_1 - 2\vec{X}_r^t\vec{X}_1\right) + \left(\vec{X}_r^t\vec{X}_r + \vec{X}_2^t\vec{X}_2 - 2\vec{X}_r^t\vec{X}_2\right) \\ + \ldots + \left(\vec{X}_r^t\vec{X}_r + \vec{X}_n^t\vec{X}_n - 2\vec{X}_r^t\vec{X}_n\right)\end{bmatrix} \\
&= -\frac{1}{2}\left[\vec{X}_r^t\vec{X}_r + \frac{1}{n}\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i - 2\vec{X}_r^t\bar{\vec{X}}\right]
\end{aligned}$$

Então:

$$-2\bar{a}_{r.} = \vec{X}_r^t\vec{X}_r + \frac{1}{n}\left(\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i\right) - 2\vec{X}_r^t\vec{\bar{X}} \tag{2.2.8}$$

Analogamente, a partir da expressão ii, pode-se mostrar que:

$$-2\bar{a}_{s.} = \vec{X}_s^t\vec{X}_s + \frac{1}{n}\left(\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i\right) - 2\vec{\bar{X}}^t\vec{X}_s \tag{2.2.9}$$

Agora, trabalhando com a expressão iii temos:

$$\bar{a}_{..} = \frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n}a_{rs}$$

$$= \frac{1}{n^2}\left[\left(a_{11}+a_{12}+...+a_{1n}\right)+\left(a_{21}+a_{22}+...+a_{2n}\right)+...+\left(a_{n1}+a_{n2}+...+a_{nn}\right)\right]$$

$$= -\frac{1}{2n^2}\left[\left(d_{11}^2+d_{12}^2+...+d_{1n}^2\right)+\left(d_{21}^2+d_{22}^2+...+d_{2n}^2\right)+...+\left(d_{n1}^2+d_{n2}^2+...+d_{nn}^2\right)\right]$$

$$= -\frac{1}{2n^2}\left\{\begin{array}{l}\left[\left(\vec{X}_1-\vec{X}_1\right)^t\left(\vec{X}_1-\vec{X}_1\right)+...+\left(\vec{X}_1-\vec{X}_n\right)^t\left(\vec{X}_1-\vec{X}_n\right)\right]+\\ +.....+\\ +\left[\left(\vec{X}_n-\vec{X}_1\right)^t\left(\vec{X}_n-\vec{X}_1\right)+...+\left(\vec{X}_n-\vec{X}_n\right)^t\left(\vec{X}_n-\vec{X}_n\right)\right]\end{array}\right\}$$

$$= -\frac{1}{2n^2}\left[\begin{array}{l}\left(n\vec{X}_1^t\vec{X}_1+\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i-2\vec{X}_1^t\sum_{i=1}^{n}\vec{X}_i\right)+\\ +.....+\\ +\left(n\vec{X}_n^t\vec{X}_n+\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i-2\vec{X}_n^t\sum_{i=1}^{n}\vec{X}_i\right)\end{array}\right]$$

$$= -\frac{1}{2n^2}\left[n\left(\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i\right)+n\left(\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i\right)-2\left(\sum_{i=1}^{n}\vec{X}_i^t\right)\left(\sum_{i=1}^{n}\vec{X}_i\right)\right]$$

$$= -\frac{1}{2n^2}\left[2n\left(\sum_{i=1}^{n}\vec{X}_i^t\vec{X}_i\right)-2\left(\sum_{i=1}^{n}\vec{X}_i^t\right)\left(\sum_{i=1}^{n}\vec{X}_i\right)\right]$$

$$= -\frac{1}{n}\left(\sum_{i=1}^{n}\vec{X}_i^t\right)+\vec{\bar{X}}^t\vec{\bar{X}}$$

Então:

$$\bar{a}_{..} = -\frac{1}{n}\left(\sum_{i=1}^{n} \vec{X}_i^t \vec{X}_i\right) + \vec{\bar{X}}^t \vec{\bar{X}} \qquad (2.2.10)$$

Substituindo as equações (2.2.8), (2.2.9) e (2.2.10) na equação (2.2.7) temos:

$$b_{rs} = \vec{X}_r^t \vec{X}_s - \vec{X}_r \vec{\bar{X}} - \vec{\bar{X}}^t \vec{X}_s + \vec{\bar{X}}^t \vec{\bar{X}}$$

Escrevendo b_{rs} na forma vetorial, temos:

$$b_{rs} = \left(\vec{X}_r - \vec{\bar{X}}\right)^t \left(\vec{X}_s - \vec{\bar{X}}\right) \qquad (2.2.11)$$

Então mostramos que cada elemento da matriz B pode ser escrito como a equação (2.2.11), além disso, a matriz B escrita em sua forma matricial é:

$$B = ZZ^t = \left(X - 1\vec{\bar{x}}\right)\left(X - 1\vec{\bar{x}}\right)$$

Logo, provamos que:

B é semi definida positiva.

Prova(⇐)

Hipótese : **B** é semi definida positiva. Então, temos que mostrar que Δ^2 é Euclidiana.
B>0 de posto "p", então, pelo Teorema de Descomposição Espectral, temos que existe uma matriz ortogonal $\Gamma=(\gamma_1,\gamma_2,\ldots,\gamma_n)$ (autovetores da matriz B) tal que:

$$\Gamma^t B \Gamma = \begin{pmatrix} [\Lambda] & [0] \\ [0] & [0] \end{pmatrix}_{nxn} \qquad (2.2.12)$$

onde:

$\Lambda=Diag(\lambda_1,\lambda_2,\ldots,\lambda_p)$ matriz diagonal e $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ são os auto-valores da matriz B.

Como a matriz **B** é de posto p<n, sabemos que os auto-valores da matriz **B**, a partir de **(p+1)** até **n** são iguais a zero, isto é:

$$\lambda_{p+1} = \lambda_{p+2} = \ldots = \lambda_n = 0$$

Logo, da equação (2.2.12) temos que:

$$B = \Gamma_{nxn} \begin{pmatrix} [\Lambda] & [0] \\ [0] & [0] \end{pmatrix}_{nxn} \Gamma^t_{nxn}$$

Então, particionando e multiplicando temos:

$$B = \Gamma_1 \Lambda \Gamma_1^t = (\Gamma_1 \Lambda^{1/2})(\Gamma_1 \Lambda^{1/2})^t = Y\, Y^t$$

onde:

$\Gamma_1 = \left(\vec{\gamma}(1), \vec{\gamma}(2), \ldots, \vec{\gamma}(p)\right)$ são os auto-vetores correspondentes aos seus respectivos auto-valores.

Escrita em sua forma matricial temos:

$$\Gamma_1 = \begin{pmatrix} \gamma_{11} & \gamma_{21} & \cdots & \gamma_{p1} \\ \gamma_{12} & \gamma_{22} & \cdots & \gamma_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1n} & \gamma_{2n} & \cdots & \gamma_{pn} \end{pmatrix}_{nxp} \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_p \end{pmatrix}_{pxp}$$

Então:

$$Y = \Gamma_1 \Lambda^{1/2}$$

isto é:

$$Y = \left[\sqrt{\lambda_1}\vec{\gamma}(1), \sqrt{\lambda_2}\vec{\gamma}(2), \ldots, \sqrt{\lambda_p}\vec{\gamma}(p)\right] = \left[\vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_p\right]$$

Já vimos que:

$$B = Y\, Y^t$$

logo, um elemento de B pode ser escrito da seguinte maneira:

$$b_{rs} = \vec{Y}_r^t \vec{Y}_s \qquad b_{rr} = \vec{Y}_r^t \vec{Y}_r \qquad b_{ss} = \vec{Y}_s^t \vec{Y}_s$$

Portanto, a distância entre $\vec{Y}_r$ e $\vec{Y}_s$ é:

$$\begin{aligned}
\left\|\vec{Y}_r - \vec{Y}_s\right\|^2 &= \left(\vec{Y}_r - \vec{Y}_s\right)^t \left(\vec{Y}_r - \vec{Y}_s\right) \\
&= \vec{Y}_r^t \vec{Y}_r + \vec{Y}_s^t \vec{Y}_s - 2\vec{Y}_r^t \vec{Y}_s \\
&= b_{rr} + b_{ss} - 2b_{rs} \\
&= \left(a_{rr} - \bar{a}_{r.} - \bar{a}_{.r} + \bar{a}_{..}\right) + \left(a_{ss} - \bar{a}_{s.} - \bar{a}_{.s} + \bar{a}_{..}\right) - 2\left(a_{rs} - \bar{a}_{r.} - \bar{a}_{.s} + \bar{a}_{..}\right) \\
&= -2a_{rs} = d_{rs}^2
\end{aligned}$$

Então:

$$\left\|\vec{Y}_r - \vec{Y}_s\right\|^2 = \left(\vec{Y}_r - \vec{Y}_s\right)^t \left(\vec{Y}_r - \vec{Y}_s\right) = d_{rs}^2$$

Por conseguinte: Δ^2 é Euclidiana. □

Finalmente, note que se T é um auto-vetor da matriz B, correspondente aos auto-valores iguais a zero então:

$$B\,T = H\,A\,H\,T = 0$$

Então T é ortogonal às colunas de Y, isto é:

$$\vec{Y}_i^t T = 0 \text{ para } i=1,2,\ldots,p$$

Analisando o vetor de médias das colunas de Y temos:

$$n\vec{\bar{Y}} = \sum_{r=1}^{p} \vec{Y}_r = Y^t T = \left[\vec{Y}_1^t T, \vec{Y}_2^t T, \ldots, \vec{Y}_p^t T\right] = \vec{0}$$

Consequentemente pode-se notar que o vetor de médias das colunas de Y é zero, isto é:

$$\vec{\bar{Y}} = 0$$

2.2.2.- Consequências do Teorema

i) Pode-se observar através da auto-estrutura da matriz **B** (onde **B** é considerada uma matriz produto escalar), que B pode ser aproximada da seguinte maneira (ver apêndice C).
Pelo Teorema do MDS métrico já provamos que **B** pode ser escrita da seguinte maneira:

$$B = ZZ^t = \left(X - 1\vec{\bar{x}}\right)\left(X - 1\vec{\bar{x}}\right)^t$$

Seber (1984) assinala que os psicometristas (Psychometricians[1]) denominam a matriz B de Q-matriz, e daí que o MDS é considerado como o método Q na Análise Multivariada (ver apêndice B).

ii) Também pelo mesmo teorema, **B** pode ser escrito em função de seus auto-valores e auto-vetores. Assim:

$$B = \Gamma_1 \Lambda \Gamma_1^t = (\Gamma_1 \Lambda^{1/2})(\Gamma_1 \Lambda^{1/2})^t = Y\,Y^t$$

onde:

$\Gamma = \left(\vec{\gamma}(1), \vec{\gamma}(2), \ldots, \vec{\gamma}(p)\right)$ matriz de auto-vetores

$\Lambda = \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ matriz diagonal de auto-valores

isto é:

$$Y = \left[\sqrt{\lambda_1}\vec{\gamma}(1), \sqrt{\lambda_2}\vec{\gamma}(2), \ldots, \sqrt{\lambda_p}\vec{\gamma}(P)\right] = \left[\vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_p\right]$$

já vimos que:

$$B = Y\,Y^t$$

Pode-se observar que a matriz **Y** está gerando uma nova configuração de pontos, de dimensão p, então uma aproximação da matriz Y é :

$$Y_{(k)} = \left[\sqrt{\lambda_1}\vec{\gamma}(1), \sqrt{\lambda_2}\vec{\gamma}(2), \ldots, \sqrt{\lambda_k}\vec{\gamma}(k)\right] = \left[\vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_k\right], \ \forall\, k<p$$

que pode ser descrito na forma matricial como:

[1] Psychometricians.- Pesquisadores que publicam na revista Psychometrika.

$$Y_{[k]} = \begin{pmatrix} \sqrt{\lambda_1}\gamma_{11} & \sqrt{\lambda_1}\gamma_{21} & \cdots & \sqrt{\lambda_k}\gamma_{k1} \\ \sqrt{\lambda_1}\gamma_{12} & \sqrt{\lambda_1}\gamma_{22} & \cdots & \sqrt{\lambda_k}\gamma_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\lambda_1}\gamma_{1n} & \sqrt{\lambda_1}\gamma_{2n} & \cdots & \sqrt{\lambda_k}\gamma_{kn} \end{pmatrix}_{nxk}$$

Isto é, estamos pegando os primeiros k auto-vetores com seus correspondentes k auto-valores (k<p). É importante notar que o melhor dos casos será quando K=2. Então a matriz calculada $Y_{[k]}$ proporciona uma nova configuração de pontos em menor dimensão, tal que as distâncias Euclidianas entre as linhas da matriz $Y_{[k]}$ sejam aproximadamente iguais às distâncias Euclidianas originais, sujeito a um erro mínimo. Muitas vezes a matriz $Y_{[k]}$ é conhecida como matriz de Coordenadas Principais, denominada assim por Gower (1966).

2.2.3.- Algoritmo prático

I] A partir da matriz de distância Δ^2 construir a matriz

$$A=[a_{rs}] \quad \text{onde:} \quad a_{rs}=(-1/2)d_{rs}^2$$

II] Obter a matriz B cujos elementos são :

$$b_{rs} = a_{rs} - \bar{a}_{r.} - \bar{a}_{s.} + \bar{a}_{..}$$

III] Achar os k maiores auto-valores $\lambda_1>\lambda_2>,...,>\lambda_k$ com seus respectivos auto-vetores padronizados: $\vec{\gamma}(1),\vec{\gamma}(2),...,\vec{\gamma}(k)$ e construir $Y_{[k]}$:

$$Y_{(k)} = \left[\sqrt{\lambda_1}\vec{\gamma}(1),\sqrt{\lambda_2}\vec{\gamma}(2),...,\sqrt{\lambda_k}\vec{\gamma}(k)\right] = \left[\vec{Y}_1,\vec{Y}_2,...,\vec{Y}_k\right]$$

O algoritmo descrito acima é um método que, basicamente fornece uma representação geométrica através das distâncias ou proximidades dos dados entre indivíduos (se conhecemos a matriz de dados). Entretanto, para confirmar este objetivo, é preciso reduzir a dimensão. Sibson (1978) propôs uma análise do erro aproximado para o procedimento MDS métrico

baseado na teoria de perturbação de Análise de Procrustes[2], e assinala que o MDS métrico é uma solução surpreendente, afetado só por um pequeno erro aleatório.

Às vezes a matriz de distância não é Euclidiana, por esta razão os auto-valores de B podem ser negativos. No entanto, como os primeiros **k** auto-valores são comparativamente maiores e positivos, os demais auto-valores podem ser positivos ou negativos, mas são próximos de zeros. Então as linhas de $Y_{[k]}$ proporcionam uma configuração razoável em **k** dimensões.

2.2.4.- Propriedades Ótimas do Escalonamento Multidimensional Métrico

A seguinte propriedade funciona supondo que a matriz de distâncias tenha sido obtida através de uma matriz de dados. Vimos que dada uma matriz de distâncias Δ^2, o objetivo do MDS métrico é encontrar uma configuração $\hat{X}$ de menor dimensão (R^k), de tal maneira que a distância entre os pontos ajustados $\hat{d}_{rs}^2 = \left(\hat{\tilde{X}}_r - \hat{\tilde{X}}_s\right)^t \left(\hat{\tilde{X}}_r - \hat{\tilde{X}}_s\right)$ seja aproximadamente igual a Δ. A partir de uma configuração ajustada $\hat{X}$ pode-se achar uma matriz de distâncias ajustadas $\hat{\Delta}$, e uma matriz de produto interno centrada ajustada $\hat{B}$ (conhecida também com o nome de matriz de dupla centralização ajustada).

Seja **X**, uma configuração inicial em $\mathbf{R}^p$ (matriz de dados) e seja $\mathbf{L}=(\mathbf{L}_1,\mathbf{L}_2)$ uma matriz ortogonal de ordem (pxp) com $\mathbf{L}_1$ de ordem (pxk). Então $\mathbf{XL}_1$ representa a projeção da configuração **X** sobre o sub-espaço de $\mathbf{R}^k$ gerado pelas colunas de $\mathbf{L}_1$. Na forma matricial, **X** e **L** são representados da seguinte maneira:
X é representado como uma matriz de dados (ver equação 2.2.1), e

$$L = [L_1, L_2] = \begin{pmatrix} l_{11} & l_{21} & \cdots & l_{k1} & l_{k+1,1} & l_{k+2,1} & \cdots & l_{p1} \\ l_{12} & l_{22} & \cdots & l_{k2} & l_{k+1,2} & l_{k+2,2} & \cdots & l_{p2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{1p} & l_{2p} & \cdots & l_{kp} & l_{k+1,p} & l_{k+2,p} & \cdots & l_{pp} \end{pmatrix}_{pxp}$$

Então pode-se achar $\hat{X} = XL_1$ como a configuração ajustada em k dimensões.

Por outro lado, como **L** é ortogonal, as distâncias entre as linhas de X são a mesma que entre as linhas de XL (ver apêndice A), isto é:

[2] Análise de Procrustes.- É uma técnica multivariada, que mede a qualidade de ajuste de duas configurações ajustadas.

$$d_{rs}^2 = \sum_{i=1}^{p} (x_{ri} - x_{si})^2 = \sum_{i=1}^{p} \left(\vec{X}_r^t l_{(i)} - \vec{X}_s^t l_{(i)} \right)^2$$

Se denotarmos as distâncias de XL_1 por $\hat{\Delta}$, então temos o seguinte:

$$\hat{d}_{rs}^2 = \sum_{i=1}^{k} \left(\vec{X}_r^t l_{(i)} - \vec{X}_s^t l_{(i)} \right)^2$$

que é a distância reduzida entre pontos da configuração projetada.
Note que: $\hat{d}_{rs}^2 \leq d_{rs}^2$

Daí, a medida de discrepância entre a configuração original **X** e a configuração projetada $\hat{X}$ é dado por:

$$\Phi = \sum_{r=1}^{n} \sum_{s=1}^{n} \left(d_{rs}^2 - \hat{d}_{rs}^2 \right)$$

Então, a solução do MDS métrico em K dimensões tem a seguinte propriedade.

Teorema 2.2.4.1

Dado Δ, a matriz de distâncias Euclidianas correspondente à configuração **X** em $\mathbf{R}^p$, e **k** fixo ($k \leq p$), então dentre todas as projeções $\mathbf{XL_1}$ de **X** sobre um sub espaço **K** dimensional de $\mathbf{R}^p$, a quantidade:

$$\Phi = \sum_{r=1}^{n} \sum_{s=1}^{n} \left(d_{rs}^2 - \hat{d}_{rs}^2 \right)$$

é minimizada quando **X** é projetado sobre a solução do MDS métrico em k dimensões.

Prova

$$\begin{aligned}
\Phi &= \sum_{r=1}^{n} \sum_{s=1}^{n} \left(d_{rs}^2 - \hat{d}_{rs}^2 \right) \\
&= \sum_{r=1}^{n} \sum_{s=1}^{n} \left[\sum_{i=1}^{p} \left(\vec{X}_r^t l_{(i)} - \vec{X}_s^t l_{(i)} \right)^2 - \sum_{i=1}^{k} \left(\vec{X}_r^t l_{(i)} - \vec{X}_s^t l_{(i)} \right)^2 \right] \\
&= \sum_{r=1}^{n} \sum_{s=1}^{n} \left[\sum_{i=k+1}^{p} \left(\vec{X}_r^t l_{(i)} - \vec{X}_s^t l_{(i)} \right)^2 \right]
\end{aligned}$$

$$= \sum_{r=1}^{n}\sum_{s=1}^{n}\left[\sum_{i=k+1}^{p} l_{(i)}^{t}\left(\vec{X}_r - \vec{X}_s\right)\left(\vec{X}_r - \vec{X}_s\right)^{t} l_{(i)}\right]$$

$$= \sum_{i=k+1}^{p}\left[l_{(i)}^{t}\left(\sum_{r=1}^{n}\sum_{s=1}^{n}\left(\vec{X}_r - \vec{X}_s\right)\left(\vec{X}_r - \vec{X}_s\right)^{t}\right) l_{(i)}\right]$$

sabe-se que : $\sum_{i=1}^{n} \vec{X}_i^{t} A \vec{X}_i = \mathrm{Tr}(AT)$ onde : $T = \sum_{i=1}^{n} \vec{X}_i \vec{X}_i^{t}$

Consequentemente temos que:

$$\Phi = \mathrm{Tr}\left[\left(\sum_{r=1}^{n}\sum_{s=1}^{n}\left(\vec{X}_r - \vec{X}_s\right)\left(\vec{X}_r - \vec{X}_s\right)^{t}\right)\left(\sum_{i=k+1}^{p} l_{(i)} l_{(i)}^{t}\right)\right]$$

$$= \mathrm{Tr}\left[\left(\sum_{r=1}^{n}\sum_{s=1}^{n}\left(\vec{X}_r - \vec{X}_s\right)\left(\vec{X}_r - \vec{X}_s\right)^{t}\right)\left(L_2 L_2^{t}\right)\right]$$

$$= \mathrm{Tr}\left\{L_2^{t}\left[\sum_{r=1}^{n}\sum_{s=1}^{n}\left(\vec{X}_r - \vec{X}_s\right)\left(\vec{X}_r - \vec{X}_s\right)^{t}\right] L_2\right\}$$

pode-se escrever : $\sum_{r=1}^{n}\sum_{s=1}^{n}\left(\vec{X}_r - \vec{X}_s\right)\left(\vec{X}_r - \vec{X}_s\right)^{t} = 2n\sum_{r=1}^{n}\left(\vec{X}_r - \vec{\bar{X}}\right)\left(\vec{X}_r - \vec{\bar{X}}\right)^{t} = 2n^2 S$

Então temos que:

$$\Phi = \mathrm{Tr}\left\{L_2^{t}\left[2n^2 S\right] L_2\right\}$$

$$= 2n.\mathrm{Tr}\left[L_2^{t} n S L_2\right]$$

Por outro lado usando a descomposição Espectral temos:

$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ autovalores de nS com autovetores normalizados Γ,

onde : $\Gamma = (\gamma_1, \gamma_2, \ldots, \gamma_p)$

Por conseguinte temos:

$$\Phi = 2n.\mathrm{Tr}\left[L_2^{t}\left(\Gamma \Lambda \Gamma^{t}\right) L_2\right]$$

$$= 2n.\mathrm{Tr}\left[L_2^{t} \Gamma \Lambda \Gamma^{t} L_2\right]$$

$$= 2n.\mathrm{Tr}\left[F_2^{t} \Lambda F_2\right]$$

$$= 2n\left[\sum_{i=1}^{p} f_{i1}^2 l_i + \sum_{i=1}^{p} f_{i2}^2 l_i + \ldots + \sum_{i=1}^{p} f_{i(p-k)}^2 l_i\right]$$

$$= 2n\left[\sum_{i=1}^{p}\sum_{j=1}^{p-k} f_{ij}^2 l_i\right]$$

$$= 2n\sum_{i=1}^{p} h_i l_i \quad \textit{onde}: \; h_i = \sum_{j=1}^{p-k} f_{ij}^2$$

Pode-se provar que:

$$\text{i) } 0 \le h_i \le 1 \qquad \text{e} \qquad \text{ii) } \sum_{i=1}^{p} h_i = p - k$$

como: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$, então $Tr[F_2^t \Lambda F_2]$ é minimizado quando: $h_1=h_2=\ldots=h_k=0$ **e** $h_{k+1}=h_{k+2}=\ldots=h_p=1$. Também observa-se que:

$$h_{k+1} = \sum_{j=1}^{p-k} f_{k+1,j}^2 = 1, \ldots, h_p = \sum_{j=1}^{p-k} f_{pj}^2 = 1$$

Logo, a forma matricial é minimizada quando $F_2 = [0, I_{p-k}]$ e, isto acontece quando $L_2=[\gamma_{(k+1)},\gamma_{(k+2)},\ldots,\gamma_{(p)}]$. Isto implica que as colunas de L_1 explicam o espaço dos primeiros k autovetores de nS, e XL_1 representa as **Coordenadas Principais** de X em k dimensões.

Note que para esta coordenada principal projetada temos:

$$\Phi = 2n[\lambda_{k+1}+\lambda_{k+2}+\ldots+\lambda_p]$$

2.2.5.- Bondade do ajuste

Uma **Medida de Ajuste**, baseada na qualidade de aproximação de matrizes (ver apêndice C), para a proporção explicada da matriz de distâncias Δ^2 pela solução clássica do MDS k-dimensional dada por Mardia, Kent e Bibby (1979), é:

$$\alpha_{1k} = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} |\lambda_i|} \text{x} 100\%$$

2.3.- MDS Métrico a partir de uma Matriz de Similaridades

Muitas vezes existem situações em que as distâncias entre n objetos, podem ser medidas através de similaridades ou preferências entre objetos por meio dos indivíduos. Então, surge a necessidade de estudar a estrutura desses tipos de matrizes.

2.3.1.- Análise de uma Matriz de Similaridades

Gower (1966) mostra que o MDS métrico, mencionado na seção 2.2, pode ser extendido para uma matriz simétrica de similaridades **S** de ordem nxn com elementos da forma s_{ij}. Mas a dificuldade de seu argumento é que Gower assume que a matriz de similaridades **S** é semi definida positiva (**S≥0**). Então, isto implica que $s_{ii}+s_{jj}-2s_{ij}$ é a verdadeira distância quadrada, ou seja;

$$\delta_{ij}^2 = s_{ii} + s_{jj} - 2s_{ij} \tag{2.3.1}$$

Agora, aplicando a transformação de Dupla Centralização sobre a matriz Δ^2, cujos elementos são dados em (2.3.1), a matriz produto escalar C como um caso geral, é expressa da seguinte maneira:

$$C = \theta S \theta t \tag{2.3.2}$$

onde θ é definida como:

$$\theta = I - \mathbf{1}(\mathbf{1}^t w)^{-1} w t \tag{2.3.3}$$

com a restrição $\mathbf{1}^t\mathbf{w}=1$

onde cada elemento da matriz C pode ser escrito da seguinte maneira:

$$c_{ij} = s_{ij} - \bar{s}_{i.} - \bar{s}_{.j} + \bar{s}_{..} \tag{2.3.4}$$

e além disso,

$$\bar{s}_{i.} = \frac{\sum_{j=1}^{n} s_{ij} w_j}{\sum_{k=1}^{n} w_k} \tag{2.3.5}$$

$$\bar{s}_{.j} = \frac{\sum_{i=1}^{n} s_{ij} w_i}{\sum_{k=1}^{n} w_k} \tag{2.3.6}$$

$$\bar{s}_{..} = \frac{\sum_{j=1}^{n}\sum_{i=1}^{n} w_j s_{ij} w_i}{\left(\sum_{k=1}^{n} w_k\right)^2} \tag{2.3.7}$$

A equação (2.3.1) é denominada por Mardia, Kent e Bibby (1979) de Transformação Padronizada. A igualdade das expressões (2.3.1) e (2.3.2) foram formalizadas no apêndice D.

A seguir mostramos um teorema, no qual provamos as afirmações de Gower (1966) como um caso particular quando θ=H (ver equação 2.3.3); onde ele afirma que se temos uma matriz de similaridades S semi definida positiva, então a matriz de distância Δ^2, definida como na equação (2.3.1), é Euclidiana, com a matriz de dupla centralização dada por B=HSH. Isto em termos práticos quer dizer que, se temos uma matriz de similaridades $S \geq 0$, podemos encontrar uma configuração de pontos através da auto-estrutura da matriz B=HSH.

2.3.2.-Teorema para uma Matriz de Similaridades.

Teorema 2.3.2.1.

Se $S \geq 0$, então a matriz de distâncias Δ^2 definida pela Transformação Padronizada (2.3.1), é Euclidiana com a matriz de produto interno centrado, B=HSH.

Prova

Por hipótese **S≥0** e pela equação (2.3.1) temos:

$$d_{ij} = (s_{ii} - 2s_{ij} + s_{jj})^{1/2}$$

e seja a matriz de similaridades:

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix}_{nxn}$$

Definamos um vetor com +1 no r-ésimo lugar e -1 no s-ésimo lugar:

$$X^t = [0,0,\ldots,+1,\ldots,-1,\ldots,0]_{1xn}$$

então podemos escrever a distância da seguinte maneira:

$$d_{ij}^{1/2} = (s_{ii} - 2s_{ij} + s_{jj}) = X^t S X \geq 0 \text{ (positiva por ser uma forma quadrática)}$$

logo, observa-se que a transformação padronizada está bem definida e Δ^2 é uma matriz de distâncias.

Sejam as matrizes **A** e **B** definidas da mesma forma que as equações (2.2.5) e (2.2.6). Vemos facilmente que HSH é uma matriz semi definida positiva, pois por hipótese S>0. Logo, **HSH>0**. Por outro lado, para provar que Δ^2 é Euclidiana é suficiente provar que **B=HSH** (ver teorema 2.2.1.1. do MDS métrico).

Já vimos que: $a_{rs} = (-1/2)d_{rs}^2$ e que se definirmos **A** e **B** como nas equações (2.2.5) e (2.2.6), temos o seguinte:

$$b_{rs} = a_{rs} - \bar{a}_{r.} - \bar{a}_{.s} + \bar{a}_{..}$$

Então:

$$\begin{aligned} -2b_{rs} &= -2a_{rs} + 2\bar{a}_{r.} + 2\bar{a}_{.s} - 2\bar{a}_{..} \\ &= d_{rs}^2 + 2\left(\frac{1}{n}\sum_{s=1}^{n} a_{rs}\right) + 2\left(\frac{1}{n}\sum_{r=1}^{n} a_{rs}\right) - 2\left(\frac{1}{n^2}\sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}\right) \end{aligned}$$

$$= d_{rs}^2 + \left(-\frac{1}{n}\right)\sum_{s=1}^{n}(-2a_{rs}) + \left(-\frac{1}{n}\right)\sum_{r=1}^{n}(-2a_{rs}) + \left(\frac{1}{n^2}\right)\sum_{r=1}^{n}\sum_{s=1}^{n}(-2a_{rs})$$

$$= d_{rs}^2 + \left(-\frac{1}{n}\right)\sum_{s=1}^{n} d_{rs}^2 + \left(-\frac{1}{n}\right)\sum_{r=1}^{n} d_{rs}^2 + \left(\frac{1}{n^2}\right)\sum_{r=1}^{n}\sum_{s=1}^{n} d_{rs}^2$$

$$= d_{rs}^2 + \left(-\frac{1}{n}\right)\sum_{i=1}^{n} d_{ri}^2 + \left(-\frac{1}{n}\right)\sum_{j=1}^{n} d_{js}^2 + \left(\frac{1}{n^2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^2$$

$$= \left[\begin{array}{l} (s_{rr} - 2s_{rs} + s_{ss}) + \left(-\frac{1}{n}\right)\sum_{i=1}^{n}(s_{rr} - 2s_{ri} + s_{ii}) + \\ + \left(-\frac{1}{n}\right)\sum_{j=1}^{n}(s_{jj} - 2s_{js} + s_{ss}) + \left(\frac{1}{n^2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}(s_{ii} - 2s_{ij} + s_{jj}) \end{array}\right]$$

$$= \left[\begin{array}{l} (s_{rr} - 2s_{rs} + s_{ss}) + \left(-s_{rr} + \frac{2}{n}\sum_{i=1}^{n} s_{ri} - \frac{1}{n}\sum_{i=1}^{n} s_{ii}\right) + \\ + \left(-\frac{1}{n}\sum_{j=1}^{n} s_{jj} + \frac{2}{n}\sum_{j=1}^{n} s_{js} - s_{ss}\right) + \left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ii} - \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_{jj}\right) \end{array}\right]$$

$$= -2s_{rs} + \left(\frac{2}{n}\right)\sum_{i=1}^{n} s_{ri} + \left(\frac{2}{n}\right)\sum_{j=1}^{n} s_{js} - \left(\frac{2}{n^2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}$$

$$= -2s_{rs} + 2\left(\frac{\sum_{i=1}^{n} s_{ri}}{n}\right) + 2\left(\frac{\sum_{j=1}^{n} s_{js}}{n}\right) - 2\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}}{n^2}\right)$$

$$= -2s_{rs} + 2\left(\frac{\sum_{s=1}^{n} s_{rs}}{n}\right) + 2\left(\frac{\sum_{r=1}^{n} s_{rs}}{n}\right) - 2\left(\frac{\sum_{s=1}^{n}\sum_{r=1}^{n} s_{sr}}{n^2}\right)$$

$$= -2s_{rs} + 2\bar{s}_{r.} + 2\bar{s}_{s} - 2\bar{s}_{..}$$

Por conseguinte, temos mostrado que:

$$b_{rs} = s_{rs} - \bar{s}_{r.} - \bar{s}_{s} + \bar{s}_{..} \qquad (2.3.8)$$

Isto implica que um elemento da matriz B=HSH pode ser escrito como na equação (2.3.8). Isto também foi visto como um caso geral na equação (2.3.4).

Logo, temos que B = H S H, implicando que a transformação padronizada é Euclidiana.

Observação

De fato é importante notar que se $S \geq 0$, então estamos assumindo que a matriz de similaridades é semi definida positiva. Por outro lado, lembremos que a matriz de projeção H é um caso particular da matriz θ (ver equação 2.3.3), onde $w=[1/n, 1/n, ..., 1/n]^t$.

2.3.3.- Consequências do Teorema

i) Do teorema acima vimos que, a partir de uma matriz de similaridades $S>0$, é possível encontrar uma configuração de pontos através da auto-estrutura da matriz de dupla centralização B=HSH, isto é, a matriz B pode ser escrita em função de seus auto-valores e auto-vetores:

$$B = \Gamma_1 \Lambda \Gamma_1^t = (\Gamma_1 \Lambda^{1/2})(\Gamma_1 \Lambda^{1/2})^t = Y\,Y^t$$

onde:

$\Gamma = \left(\vec{\gamma}(1), \vec{\gamma}(2), \ldots, \vec{\gamma}(p)\right)$ matriz de auto-vetores.

$\Lambda = \mathrm{Diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ matriz diagonal de auto-valores.

isto é:

$$Y = \left[\sqrt{\lambda_1}\vec{\gamma}(1), \sqrt{\lambda_2}\vec{\gamma}(2), \ldots, \sqrt{\lambda_p}\vec{\gamma}(p)\right] = \left[\vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_p\right]$$

Uma aproximação da matriz Y é:

$$Y = \left[\sqrt{\lambda_1}\vec{\gamma}(1), \sqrt{\lambda_2}\vec{\gamma}(2), \ldots, \sqrt{\lambda_p}\vec{\gamma}(k)\right] = \left[\vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_k\right]$$

que, escrito na forma matricial, fica:

$$Y_{[k]} = \begin{pmatrix} \sqrt{\lambda_1}\gamma_{11} & \sqrt{\lambda_1}\gamma_{21} & \cdots & \sqrt{\lambda_k}\gamma_{k1} \\ \sqrt{\lambda_1}\gamma_{12} & \sqrt{\lambda_1}\gamma_{22} & \cdots & \sqrt{\lambda_k}\gamma_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\lambda_1}\gamma_{1n} & \sqrt{\lambda_1}\gamma_{2n} & \cdots & \sqrt{\lambda_k}\gamma_{kn} \end{pmatrix}_{nxk}$$

Isto é, estamos pegando os primeiros k autovetores com seus correspondentes k auto-valores (k<p). É importante notar que o melhor dos casos será quando K=2. Esta configuração de pontos obtida através de B indica a representação geométrica dos n indivíduos por n pontos, a qual foi denominada de Coordenadas Principais por Gower (1966). Como se poderá observar, a matriz de similaridades S pode ser comparada com matriz A definida na equação (2.2.5).

ii) A técnica também pode ser aplicada a qualquer matriz de distância não necessariamente Euclidiana.

2.3.4.- Bondade do ajuste

Neste caso, Mardia, Kent e Bibby (1979) propuseram uma medida de discrepância quando estamos trabalhando com uma matriz de similaridades ou uma matriz de distâncias, não necessariamente Euclidiana. Se $\hat{X}$ é uma nova configuração ajustada através da auto-estrutura da matriz B, com matriz de produto interno centrado ajustada $\hat{B}$, então a medida de discrepância entre B e $\hat{B}$, é dada por:

$$\Psi = \sum_{r=1}^{n}\sum_{s=1}^{n}\left(b_{rs} - \hat{b}_{rs}\right)^2 = Tr\left(B - \hat{B}\right)^2$$

Teorema 2.3.4.1

Se Δ^2 é matriz de distâncias (não necessariamente Euclidiana), então para um k fixo $\Psi = \sum_{r=1}^{n}\sum_{s=1}^{n}\left(b_{rs} - \hat{b}_{rs}\right)^2 = Tr\left(B - \hat{B}\right)^2$ é minimizado sobre qualquer configuração $\hat{X}$ em k dimensões quando $\hat{X}$ é a solução do problema MDS métrico.

Neste caso a medida de ajuste é:

$$\alpha_{2k} = \frac{\sum_{i=1}^{k} \lambda_i^2}{\sum_{i=1}^{n} \lambda_i^2} x100\%$$

Pode-se observar que esta medida está baseada na aproximação de matrizes vista na parte de estrutura de matrizes (apêndice B, ver as equações c14 e c15).

2.4.- MDS Métrico de Quadrados Mínimos

Vimos que o MDS métrico tenta achar uma configuração de pontos, de tal forma que as distâncias estimadas entre estes pontos seja aproximadamente igual às distâncias (ou similaridades) observadas. Tudo isto foi feito nas seções anteriores deste capítulo, usando basicamente teoria algébrica de aproximação de matrizes. Há outras formas de trabalhar, como otimizar medidas de discrepância. Krzanowski e Marriott (1994) assinalam que Sammon (1969) apresenta um trabalho no qual ele estima uma configuração de pontos em k dimensões, através da minimização da medida de discrepância dada por:

$$y = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\delta_{ij} - d_{ij}\right)^2 \qquad (2.5.1)$$

onde:

δ_{ij}: distância ou similaridade observada.

d_{ij}: distância calculada através de uma configuração de pontos.

Como se poderá observar, uma solução algébrica não é possível, por isso recorre-se a métodos numéricos. A configuração inicial é especificada escolhendo o número de dimensões k, e a medida de discrepância é vista então como uma função de nk variáveis, sendo minimizada usando o método de Steepest Descent [3]. A solução do MDS métrico em k-dimensões pode ser considerada como uma configuração inicial. O procedimento continua com processos iterativos até encontar o mínimo valor da medida de discrepância em (2.5.1).

Mais geralmente para o processo de ajuste e, ao mesmo tempo para os aspectos ótimos de locação e escala, as distâncias ajustadas d_{ij} podem ser regressadas sobre as distâncias (ou

[3] Steepest Descent.- Método numérico de minimização de funções com várias variáveis. (f: $\mathbb{R}^n \rightarrow \mathbb{R}$).

similaridades) observadas δ_{ij}. Esta análise vai produzir um conjunto de distâncias estimadas ($\hat{d}_{ij}$), as quais são muitas vezes chamadas de disparidades, isto é:

$$\hat{d}_{ij} = \hat{a} + \hat{b}\delta_{ij}$$

onde $\hat{a}$ e $\hat{b}$ são parâmetros estimados através da Análise de Regressão. A soma de quadrados dos resíduos obtida através desta Regressão fornece outras medidas de discrepância. Para comparabilidade através de situações diferentes, esta soma de quadrados dos resíduos pode ser normalizada por: $\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^2$, produzindo uma soma de quadrados dos resíduos padronizadas, definida por:

$$\text{STRESS} = \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(d_{ij} - \hat{d}_{ij}\right)^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^2} \right)^{\frac{1}{2}}$$

que foi introduzida, pela primeira vez por Kruskal (1964 a, b), e que será melhor analisado na parte de MDS não métrico (capítulo III).

Por outro lado, De Leeuw, Young e Takane (1976), usam quadrados mínimos alternados[4] em vez do método de steepest descent para o processo de minimização e sugerem substituir d_{ij} e $\hat{d}_{ij}$ por seus quadrados, produzindo a seguinte medida de discrepância denominada de S-STRESS:

$$\text{S-STRESS} = \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(d_{ij}^2 - \hat{d}_{ij}^2\right)^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^4} \right)^{\frac{1}{2}}$$

Detalhes desta análise são vistos no capítulo III, na parte da descrição do procedimento ALSCAL.

[4] Quadrados mínimos alternados.- Um procedimento numérico para minimizar funções, que nesta parte métrica do MDS é usado como um problema de constante aditiva. Maiores detalhes ver Young, DeLeeuw e Takane (1976).

Capítulo III

OUTROS TIPOS DE ESCALONAMENTO MULTIDIMENSIONAL

3.1.- Introdução

Este capítulo apresenta outras formas de analisar a técnica de Escalonamento Multidimensional (MDS). Descrevemos, basicamente o problema de Escalonamento Multidimensional não Métrico (MDS não métrico), técnica que é considerada complementar ao MDS métrico. O MDS não métrico é caracterizado por dois aspectos. O primeiro, é o nível de mensuração dos dados. Se eles estiverem na escala ordinal ou nominal, o MDS é não métrico. O segundo aspecto diz respeito a uma atitude mais específica do MDS não métrico em relação ao MDS Métrico, atitude esta que toma como hipótese uma relação menos rígida entre d_{ij} e δ_{ij} dada por:

$$\delta_{ij} = f\left(d_{ij} + e_{ij}\right)$$

onde **f** é uma função desconhecida monótona. Neste caso a única informação a ser usada para a reconstrução da configuração é a ordenação das δ_{ij}.

Também analisaremos o problema para o caso em que temos uma matriz de distâncias ou similaridades para cada indivíduo, isto é, cada indivíduo vai ter uma maneira diferente de dar sua preferência ou semelhança, e a partir dessa informação individual trataremos de encontrar uma configuração comum de pontos para todos os indivíduos, a qual estará indicando o comportamento dos estímulos. Davison (1983) assinala que a idéia surge com Bloxom (1968) e Horan (1969), depois é escrita por Caroll e Chang (1970) que introduzem o termo

Escalonamento de Diferenças Individuais (Individual Differences Scaling), que também é conhecido como Modelo Euclidiano Ponderado (The Weighted Euclidean Model). Depois detalha-se a análise de matrizes retangulares de dados de preferência. Na última parte descreve-se, simplificadamente, o modelo "Três Modos" (Three mode), baseado na idéia de Tucker (1972), como uma generalização do Modelo Euclidiano Ponderado.

3.2.- Escalonamento Multidimensional Não Métrico (MDS não métrico)

A análise do MDS não métrico surge nos anos 60, com o aporte principal de Coombs (1958), que desenvolveu o método de MDS não métrico baseado na ordem dos dados de proximidade, isto é, ele partia de apenas uma ordenação das medidas de proximidade. Depois Shepard (1962 a, b) elaborou um algoritmo que permitia obter uma solução métrica a partir de dados não métricos; o seu método ficou conhecido como Análise de Proximidades. O aporte de Kruskal (1964 a, b) foi introduzir uma medida de bondade do ajuste, procurando obter a configuração que melhor se ajustasse aos dados, usando métodos numéricos. Kruskal exigiu uma relação monotônica crescente ou decrescente entre as medidas de proximidade e as distâncias na configuração procurada. Supôs que há uma configuração de pontos num espaço euclidiano k-dimensional que é verdadeira, com a qual ele poderia descobrir apenas a ordenação linear das distâncias entre os pontos, podendo, a partir desta informação não métrica, recuperar a configuração. Percebe-se este fato quando apenas a informação ordinal (ordem ou posto) é usada, e desse modo obtém-se uma configuração de pontos em menor dimensão. Este processo é chamado de MDS não métrico. Greenacre e Underhill (1982) afirma que todas as idéias centrais de MDS não métrico trabalham sobre Regressão Monotônica (Monotonic Regression), também conhecida como Regressão Isotônica (Isotonic Regression), devido ao fato de Kruskal (1962) ter desenvolvido a teoria de MDS não métrico baseado na ordem das similaridades ou distâncias, para o qual introduziu uma medida de ajuste chamada de STRESS construída a partir de uma Regressão Monotônica. Os trabalhos de Levy (1981), Seber (1984), Jain e Dubes (1988), Manly (1994) e Krzanowski e Marriott (1994) também ajudam a entender o significado do MDS não métrico.

3.2.1.- Problema Geral de MDS Não Métrico

Supondo que temos uma matriz de distâncias ou similaridades observadas $\Delta=(\delta_{ij})$, podemos construir uma configuração de pontos em menor dimensão (digamos k), a qual vamos considerar como configuração inicial (existem várias maneiras de achar uma configuração inicial, as quais especificaremos mais adiante). A partir desta configuração inicial, calcularemos as distâncias d_{ij}. Agora tentaremos reconstruir a configuração original Δ de distâncias observadas, cujos elementos são da forma:

$$\delta_{ij} = d_{ij} + e_{ij} \tag{3.2.1}$$

onde os e_{ij} representam os erros de medição, devido à configuração estimada em R^k ser uma aproximação da configuração original. Shepard (1962) diz que os δ_{ij} estão relacionados de alguma maneira com os d_{ij}, os quais modificam a equação (3.2.1) da seguinte maneira:

$$\delta_{ij} = f\left(d_{ij} + e_{ij}\right) \tag{3.2.2}$$

onde **f** é uma função monótona não-crescente desconhecida. Para o modelo (3.2.2), a única informação que Shepard (1962) usa para reconstruir d_{ij} é a ordem dos postos de δ_{ij}. Como se poderá observar, a maior contribuição é o uso do critério de monotonicidade entre as similaridades ou distâncias observadas (δ_{ij}) e as distâncias calculadas (d_{ij}).

Nesta parte da aproximação não métrica, Mardia, Kent e Bibby (1979) sugerem que trabalhar sobre matrizes de similaridades é mais apropriado do que trabalhar sobre matrizes de distâncias.

3.2.2.- Solução de MDS não métrico

A solução do MDS não métrico foi desenvolvida por Shepard (1962) e Kruskal (1964) que iniciam os trabalhos com uma variedade de aproximações não métricas. A seguir descrevemos os métodos:

3.2.2.1. Método de Shepard (1962).

Descreveremos de forma resumida, a idéia principal do método de Shepard, sem nos preocuparmos com as questões numéricas que ela envolve. Shepard tinha por objetivo construir

uma configuração de pontos num espaço de menor dimensão de modo que as distâncias entre os pontos da configuração estivessem monotonicamente relacionados com as medidas de proximidade entre os estímulos. Isto é, Shepard procurava uma configuração de pontos $\vec{x}_1,\vec{x}_2,...,\vec{x}_n \in R^k$ tal que **k** fosse mínimo, e de modo que $d\left(\vec{x}_i,\vec{x}_j\right) = f\left(\delta_{ij}\right)$, com f sendo uma função monótona. Daqui em diante, denotaremos a configuração de pontos por X, que é escrita matricialmente como:

$$X_{nxk} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \begin{matrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_n \end{matrix} \qquad (3.2.3)$$

Uma vez colhidas as medidas de proximidade δ_{ij}, Shepard partia para um método iterativo, visando obter a configuração de pontos $\vec{x}_1,\vec{x}_2,...,\vec{x}_k \in R^k$ que haveria de representar os estímulos num espaço euclidiano R^k. A idéia de Shepard foi dividida em duas partes:

Primeiro : Escolha uma configuração inicial arbitrária.

Segundo : Nesta etapa, Shepard movimentava os pontos da configuração. Para fazer isto ele baseava-se em dois objetivos:
- O primeiro objetivo era diminuir o número de dimensões do espaço (Shepard iniciava a análise no espaço de maior dimensão possível).
- O segundo objetivo era procurar aumentar o ajuste dos pontos da configuração à ordenação das medidas de proximidade observadas.

3.2.2.2. Método de Kruskal (1964).

A análise de Kruskal inicia-se pela procura de uma função objetiva, digamos S. Para se definir esta função existe uma condição: a função S tem que medir o ajuste da configuração X à matriz de distâncias ou similaridades observadas. Isto é, dado uma matriz de distâncias ou similaridades observadas $\Delta=(\delta_{ij})$ de ordem nxn, calcular uma configuração de pontos X_{nxk} num espaço Euclidiano k-dimensional, que reproduza **m=n(n-1)/2** distâncias entre pontos, denotadas por $D=(d_{ij})$. A condição é que a função S deve produzir um número que mostre o

quanto a solução está ajustada aos dados. Para definir a função S, Kruskal impõe que a distância entre os pontos obtidos seja uma função monotônica das proximidades observadas:

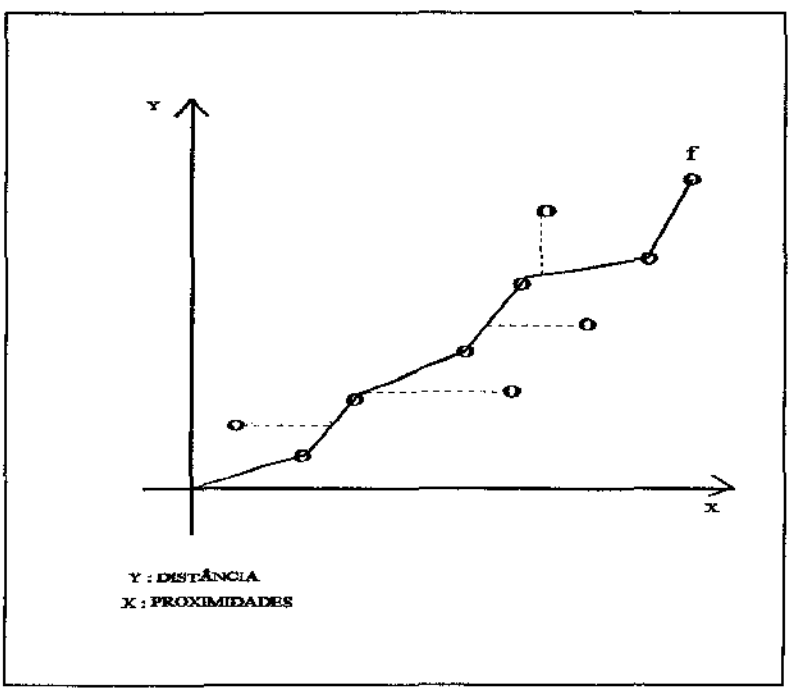
$$d_{ij} = f\left(\delta_{ij}\right) \tag{3.2.4}$$

onde f deve ser uma função monótona crescente ou decrescente. Isto é, se considerarmos uma função crescente, esta deve satisfazer o seguinte:

$$\text{se}: \ f\left(\delta_{ij}\right) \leq f\left(\delta_{kl}\right) \Rightarrow d_{ij} \leq d_{kl} \tag{3.2.5}$$

Desta maneira, o objetivo é estimar uma relação de monotonicidade não paramétrica entre Δ e D. Em geral, a equação (3.2.5) não é satisfeita para qualquer configuração X. Por essa razão, surge a necessidade de avaliar e calcular a função objetiva S, que deve produzir um número que mostre o quanto a solução está ajustada aos dados. Pensando em definir a função S, Kruskal foi o mais intuitivo possível. Começou analisando as discrepâncias existentes entre d_{ij} e $f(\delta_{ij})$, que podem ser observadas nas linhas pontilhadas, no gráfico 3.1, que mostra uma função f qualquer monotônica.

Gráfico 3.1. Gráfico de discrepância entre d_{ij} e $f(\delta_{ij})$



As discrepâncias são dadas por d_{ij}-f(δ_{ij}), não interessando o seu sinal. Tanto as positivas como as negativas, são interpretadas igualmente e consideradas indesejáveis. Kruskal então define uma função objetiva S, chamada f-STRESS, em função das discrepâncias quadráticas. O f-STRESS, em sua primeira versão foi dado pela fórmula:

$$\text{f - STRESS} = \sum_{i=1}^{m}\sum_{j=1}^{m}\left(d_{ij} - f\left(\delta_{ij}\right)\right)^2 \tag{3.2.6}$$

O f-STRESS, definida na equação (3.2.6), é invariante com respeito à rotação, reflexão e translação de X (isto quer dizer, à transformações de X para AX+B, com A matriz ortogonal e B matriz arbitrária). Para que a medida de STRESS seja invariante com respeito à dilatação, é necessário incluir o fator escala, isto é:

$$\text{f - STRESS} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left[d_{ij} - f\left(\delta_{ij}\right)\right]^2}{\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}^2} \tag{3.2.7}$$

$$\text{f - STRESS} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left(d_{ij} - f\left(\delta_{ij}\right)\right)^2}{\sum_{i=1}^{m}\sum_{j=1}^{m}\left(d_{ij} - \bar{d}_{ij}\right)^2} \tag{3.2.8}$$

onde: $\bar{d}_{ij} = \dfrac{\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}}{m}$ e as equações (3.2.7) e (3.2.8) são invariantes com respeito à dilatação (isto quer dizer, à transformações de X para bX, onde b é um escalar). Em aplicações de MDS não métrico, o uso das equações (3.2.7) ou (3.2.8) causa uma pequena diferença na prática, mas quase sempre é recomendável a equação (3.2.7). Para a análise apresentada neste trabalho, usaremos a equação (3.2.7).

Percebe-se que quanto maior for o f-STRESS, pior é o ajuste da configuração e da função f às observações. O f-STRESS, obviamente, sempre é não negativo. Se for nulo,

implica que $d_{ij}=f(\delta_{ij})$ $\forall i,j$, e tem-se, portanto, a representação perfeita, no sentido das distâncias estarem perfeitamente relacionadas com as proximidades pela função f.

Em seguida, procura-se uma medida de ajuste de configuração, independente da função f. Define-se então:

$$\text{STRESS}(\Delta,X) = \min (\text{f-STRESS}(\Delta,X)) \quad \forall f \tag{3.2.9}$$

significando que está sendo usada a melhor função f para a configuração dada. Pode-se notar que a monotonicidade exigida pelo modelo nem sempre é obtida na configuração, então definimos:

$$\hat{d}_{ij} = f(\delta_{ij}) \tag{3.2.10}$$

onde $\hat{d}_{ij}$ são números monotonicamente relacionados com as proximidades δ_{ij} (isto é, satisfazem a desigualdade 3.2.5) e minimizam o STRESS dado por:

$$S = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left[d_{ij} - \hat{d}_{ij}\right]^2}{\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}^2} = \text{STRESS}(\Delta, X) \tag{3.2.11}$$

Uma vez definido o STRESS, define-se a configuração X a partir das observações Δ. A obtenção da melhor configuração consiste na obtenção daquela que produz o menor valor possível para o STRESS. Esta configuração $\hat{X}$ é aquela que tem o menor S, para uma dada matriz de proximidades Δ. Ou seja, a configuração $\hat{X}$ é aquela para a qual é válida a igualdade:

$$\text{STRESS}(\Delta,\hat{X}) = \min(\text{STRESS}(\Delta,X)) \quad \forall X \tag{3.2.12}$$

3.2.2.3. Procedimentos para Obtenção da Solução do Método de Kruskal.

A seguir descrevemos as etapas do procedimento para obter a solução do MDS não métrico, pelo método de Kruskal.

- Primeira Etapa

Escolher uma configuração inicial X arbitrária num espaço k-dimensional. Note que quando a escolha da configuração inicial é bem feita, economiza-se várias iterações no processo de obtenção da solução do MDS não métrico. Para escolher a configuração X arbitrária deve-se obedecer duas condições:

i) $\vec{x}_i \neq \vec{x}_j \forall i,j$, onde $\vec{x}_i$ *e* $\vec{x}_j$ são pontos da configuração X.

ii) X não pode estar contida em R^t tal que t<k, onde k é a dimensão do espaço da configuração procurada.

Sugestões para escolher a configuração inicial:

a.- Iniciar com a solução do MDS métrico.
b.- Gerar os pontos $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n$ a partir de um gerador de números aleatórios.

Qualquer que seja a escolha da configuração X, esta deve satisfazer duas condições:
- X deve ser normalizado, isto é, o seu centro de gravidade deve localizar-se na origem do sistema de coordenadas.
- A soma das distâncias de seus pontos à origem do sistema deve ser igual a 1, isto é, $\sum_{i=1}^{n} \sum_{j=1}^{k} x_{ij}^2 = 1$.

- Segunda Etapa

Esta etapa refere-se a obtenção dos $\hat{d}_{ij}$. Ao se estimar os $\hat{d}_{ij}$, deve-se lembrar que estes são apenas números reais, não são distâncias. Não há uma configuração cujas distâncias entre os pontos sejam $\hat{d}_{ij}$. Os $\hat{d}_{ij}$ são apenas uma seqüência monotônica de números, escolhidos o mais próximo possível dos d_{ij}, que são usados como referência para o grau de monotonicidade das medidas d_{ij} . Logo, pode-se considerar a estimativa dos d_{ij} como um problema de Regressão Monotônica (ver apêndice E) que consiste em procurar os $\hat{d}_{ij}$ sujeitos à restrição de monotonicidade, e que tornam mínima a soma de quadrados do numerador do STRESS (ver equação 3.2.11), uma vez que, para uma dada configuração X, a soma $\sum_{i=1}^{m} \sum_{j=1}^{m} d_{ij}^2$ é fixa.

- Terceira Etapa

Esta etapa consiste apenas do cálculo do stress e da análise do seu resultado. Isto é, através dos $\hat{d}_{ij}$ é obtido o STRESS para a configuração X. Se o valor do STRESS satisfizer as condições do experimentador, esta configuração é a solução do MDS não métrico, caso contrário passa-se a etapa seguinte.

- Quarta Etapa

Procura-se uma nova configuração X, que minimize o STRESS. O problema consiste na otimização do STRESS (isto é, na minimização). Sabe-se que o STRESS definido na equação (3.2.11) é função da configuração X, pois as distâncias d_{ij} são obtidas a partir da configuração X. Isto é:

$$S = STRESS(\Delta, X) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left[d_{ij} - \hat{d}_{ij}\right]^2}{\sum_{i=1}^{m}\sum_{j=1}^{m} d_{ij}^2} = f\left(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\right) \qquad (3.2.13)$$

Logo, através da equação (3.2.13), percebe-se que o STRESS é uma função de várias variáveis, pois temos n pontos, cada ponto pertencente ao espaço k-dimensional ($\vec{x}_i \in R^k$ *onde :* $i = 1,2,\ldots,n$). Concluímos, então, que o STRESS é uma função de nk variáveis. A solução do MDS não métrico é dada pela configuração $\hat{X} = \left(\hat{\vec{x}}_1, \hat{\vec{x}}_2, \ldots, \hat{\vec{x}}_n\right)$ que minimiza a função S. Esta solução pode ser obtida resolvendo um problema de Análise Numérica que usa a técnica conhecida como Método dos Gradientes[1] ou Método Steepest-descent[2]. Retorna-se em seguida à segunda etapa do procedimento

3.2.3.- Algoritmo de Shepard-Kruskal

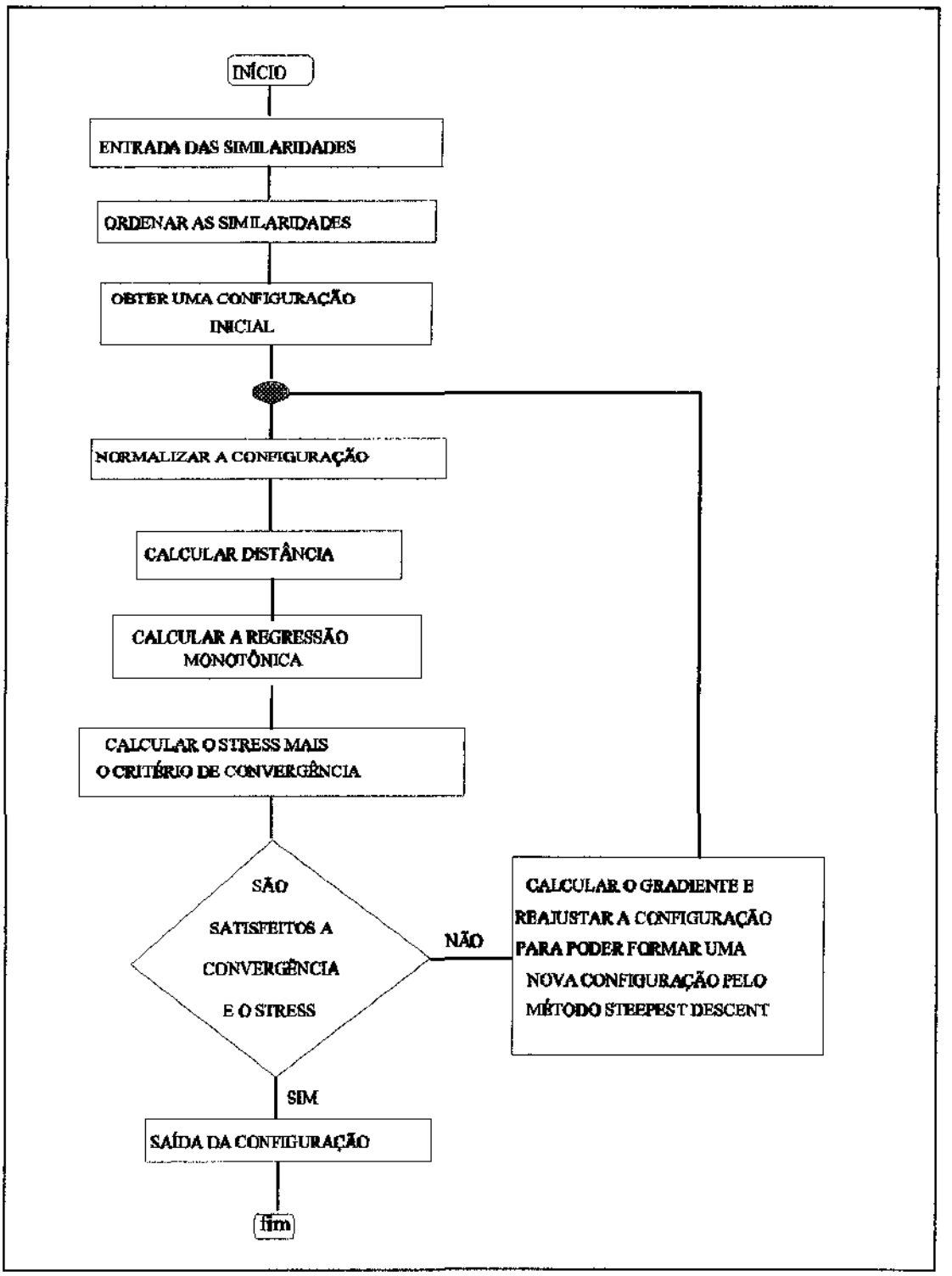
Nesta parte apresentamos o diagrama que mostra como foi desenvolvido o algoritmo prático para o MDS não métrico. A idéia é juntar os dois métodos, o de Shepard e o de Kruskal, pois segundo a literatura, o método final de Kruskal foi baseado no método de Shepard.

---

[1] Método dos Gradientes.- Método numérico, usado para minimizar funções, usam uma medida de declividade da função para indicar a direção do mínimo.

[2] Método steepest descent.- Método numérico de minimização de funções com várias variáveis. (f: $\mathbb{R}^n \rightarrow \mathbb{R}$).

Grafico 3.2. Diagrama do Algoritmo de Shepard e Kruskal



3.2.4. Regra de Kruskal.

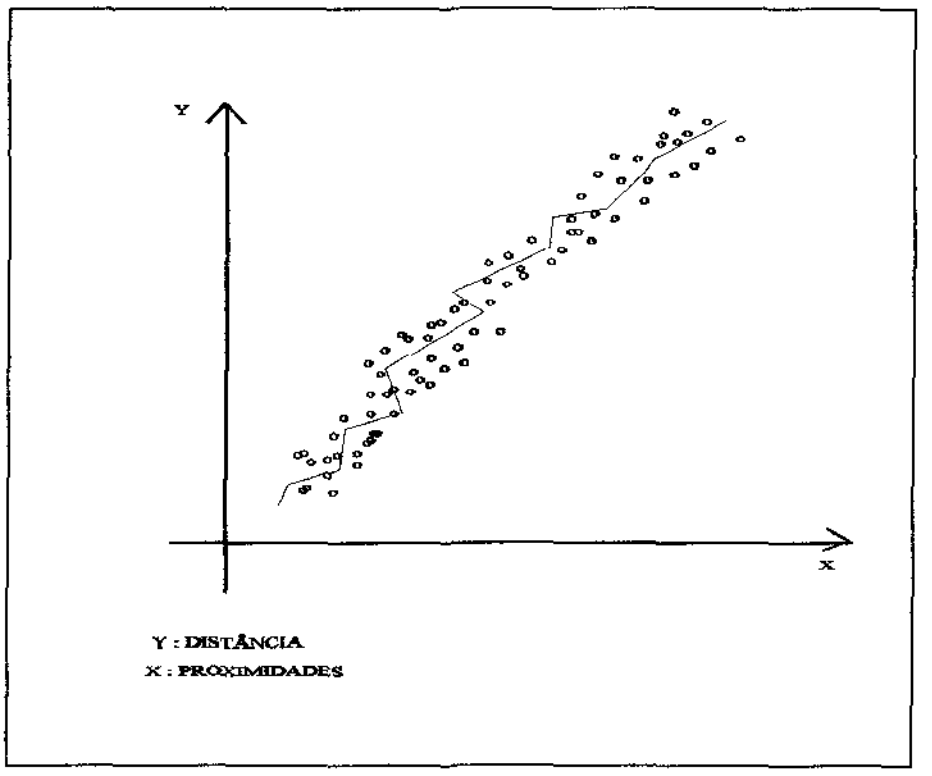
A regra dada por Kruskal (1964), para avaliar a tolerância do STRESS é a seguinte:

STRESS ≥ 20%	**INSATISFATORIO-FRACO**
5%≤ STRESS ≤ 20%	**SATISFATÓRIO-FAVORÁVEL**
STRESS ≤ 5%	**BOM**
STRESS = 0	**PERFEITO**

3.2.5. Diagrama de Dispersão

O diagrama de dispersão para uma determinada configuração é um gráfico das medidas d_{ij} em função das medidas δ_{ij}. Portanto, este diagrama tem assinalado um ponto para cada par de estímulos.

Gráfico 3.3. Diagrama de Dispersão.



Este diagrama foi utilizado na definição da função de ajuste e pode também ser um auxiliar do próprio stress na avaliação do ajuste da configuração aos dados. Uma vez obtida a solução do MDS, pode-se construir um diagrama de dispersão com duas finalidades. Primeiramente, verificar se a relação exigida entre distâncias d_{ij} e medidas de proximidade δ_{ij} foi obedecida nessa solução. Secundariamente, procurando obter uma impressão visual do ajuste da configuração aos dados, uma vez que já foi obtido o valor numérico do ajuste através do STRESS. Outra forma de verificar, seria achando o coeficiente de Correlação de Pearson entre $\hat{d}_{ij}$ e δ_{ij}, pois eles devem exibir uma dependência linear, isto é, uma correlação alta indicará um bom ajuste da configuração.

3.3.- Modelo Euclidiano Ponderado

Pode-se perceber que todos os métodos mencionados até agora sobre Escalonamento Multidimensional (MDS) são apropriados para analisar uma só matriz de distâncias ou similaridades. O Modelo Euclidiano Ponderado (The Weighted Euclidean Model), muitas vezes conhecido como Escalonamento de Diferenças Individuais (Individual Differences Scaling), é denominado assim por Caroll e Chang (1970), devido ao fato de cada indivíduo ser associado a uma matriz de distâncias ou similaridades.

Em muitas aplicações, particularmente em áreas de Psicologia e Estudo Sensorial, apresentam-se uma série de matrizes de distâncias ou similaridades para análise. Por exemplo, em uma comparação de bebidas, cada membro de um painel de provadores (ou experimentadores) pode fornecer uma matriz de similaridades entre pares de bebidas; em mercadologia, indivíduos em diferentes partes de uma cidade podem ser requisitados para comparar vários produtos que estão competindo no mercado e, em Psicologia, indivíduos de diferentes grupos de idade podem ser requisitados para julgar as similaridades entre vários pares de estímulos. Às vezes, este tipo de análise é considerado como um Escalonamento Multidimensional sobre estrutura de grupos.

A literatura aponta três aproximações comums para analisar o modelo. Bloxom (1968), Horan (1969) e Caroll e Chang (1970) descrevem o primeiro modelo incorporando um peso para cada indivíduo sobre o modelo original de distância métrica de Torgerson (MDS métrico). Este modelo é muitas vezes chamado de INDSCAL, por ser identificado com o programa de computação INDSCAL, proposto na primeira versão por Caroll e Chang (1970) e usado na solução da estimativa dos parâmetros.

Para evitar confusão entre o modelo e o programa de computação, o modelo em questão será chamado de Modelo Euclidiano Ponderado. A segunda aproximação é o

procedimento não métrico que ajusta o Modelo Euclidiano Ponderado denominado de ALSCAL, desenvolvido por Takane, Young e De Leeuw (1976), ele trabalha com uma medida de ajuste semelhante do STRESS denominado de S-TRESS. Outra aproximação mais geral é chamada de Modelo Três modos para o MDS (Three-mode Model to MDS); este modelo foi usado por Tucker (1972) em alguns estudos psicofísicos, e será descrito, de maneira resumida, na parte final do presente capítulo.

Iniciaremos a discussão do Modelo Euclidiano Ponderado, baseado na teoria de Bloxom (1974), Davison (1983), Seber (1984), Krzanowski e Marriott (1995). Ressaltamos que este modelo é muito aplicado na prática e, além disso, os programas de computação voltados para ele foram avaliados por Schiffman, Reynolds e Young (1982) e a maior parte dos principais softwares estatísticos o mencionam.

3.3.1.- Problema Geral do "Modelo Euclidiano Ponderado"

Suponha que temos **S** indivíduos e **I** estímulos, cada indivíduo com sua respectiva matriz de similaridades ou distâncias:

$$\Delta_S = \delta_{ijs} \quad i,j=1,2,...,\mathbf{I} \quad s=1,2,...,\mathbf{S}$$

Definimos a matriz de dupla centralização:

$$B_S = HA_SH = \left[I - \left(\frac{1}{n}\right)\mathbf{1}\,\mathbf{1}^t\right]A_S\left[I - \left(\frac{1}{n}\right)\mathbf{1}\,\mathbf{1}^t\right]$$

onde:

I : Matriz de identidade.

1 : Vetor de uns.

$A_i = (-1/2)\,\delta_{ijs}^2$

Então, o procedimento do MDS métrico ou não métrico produz uma configuração de pontos p-dimensional com coordenadas:

$$X_s = (x_{iks}) \quad \text{para } i=1,2...,I. \quad k=1,2...,K. \quad s=1,2,...,S.$$

Esta matriz de coordenadas estimadas também é conhecida como matriz de coordenadas dos estímulos, e satisfaz:

$$B_s = X_s X_s^t$$

onde X_S é uma matriz de coordenadas dos estímulos de ordem (IxK) para o indivíduo **s** com elementos x_{iks}. Aqui x_{iks} é a coordenada do indivíduo "**s**" para o estimulo "**i**" na dimensão "**k**". Por outro lado, consideremos uma matriz X de coordenadas dos estímulos comums para todos os indivíduos com elemento x_{ik}, onde x_{ik} indica a coordenada do estímulo **i** na dimensão **k**. Os elementos da matriz X_s (de cada indivíduo), isto é, x_{iks} são supostamente relacionados aos elementos da matriz de coordenadas dos estímulos comums X. Isto quer dizer que x_{iks} está relacionada de alguma maneira com x_{ik}.

3.3.2.- Solução do Modelo Euclidiano Ponderado

Como estamos supondo que X_s está relacionado com X, então, Caroll e Chang (1970) postulam um modelo relacionando o conjunto de coordenadas dos estímulos comums x_{ik} e pesos w_{ks} com as coordenadas de cada indivíduo x_{iks} da seguinte forma:

$$x_{iks} = x_{ik}\, w_{ks} \qquad (3.3.1)$$

onde: i=1,2,...,I k=1,2,...,K e s=1,2,...,S

Se W_s é definido como uma matriz diagonal de ordem (kxk), isto é:

$$W_s = \mathrm{Diag}\left[w_{s1}, w_{s2}, \ldots, w_{sK}\right] = \begin{pmatrix} w_{s1} & 0 & \cdots & 0 \\ 0 & w_{s2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{sK} \end{pmatrix}$$

então, a equação (3.3.1) é escrita na forma matricial da seguinte maneira:

$$X_s = X\, W_s \qquad (3.3.2)$$

Da equação (3.3.2) podemos perceber que, a partir da configuração para cada indivíduo X_s vamos tentar encontrar uma só matriz de coordenadas dos estímulos comums X para todos os indivíduos e uma matriz de peso W_s respectivo a cada indivíduo. De acordo com o modelo (3.3.1) podemos calcular a distância Euclidiana entre os estímulos a partir da configuração X_s, isto é,

$$d_{ijs} = \left[\sum_{k=1}^{K} \left(x_{iks} - x_{jks} \right)^2 \right]^{\frac{1}{2}} \tag{3.3.3}$$

Na equação (3.3.3), d_{ijs} refere-se à distância entre os estímulos i e j a partir da configuração do indivíduo s. Substituindo a relação (3.3.1) na equação (3.3.3) temos:

$$d_{ijs} = \left[\sum_{k=1}^{K} w_{ks}^2 \left(x_{ik} - x_{jk} \right)^2 \right]^{\frac{1}{2}} \tag{3.3.4}$$

A equação (3.3.4) expressa a suposição fundamental do Modelo Euclidiano Ponderado. Estes pesos w_{ks} são às vezes chamados pesos importantes. Também pode-se observar que se os pesos aumentam, haverá influências sobre as distâncias entre os estímulos i e j.

3.3.3.- Ajustando o Modelo Euclidiano Ponderado

Nesta seção daremos uma idéia de como ajustar os parâmetros do Modelo Euclidiano Ponderado. Da equação (3.3.2) temos :

$$X_s = X\, W_s$$

então o problema é encontrar as coordenadas da matriz comum para todos os indivíduos X e os pesos W_s para cada indivíduo.

Por outro lado, como conhecemos as matrizes de similaridades ou distâncias para cada indivíduo, podemos pelo MDS métrico, encontrar uma configuração de pontos tal que:

$$B_s = X_s X_s^t \tag{3.3.5}$$

Substituindo-se a equação (3.3.2) na equação (3.3.5) temos:

$$B_S = XW_S W_S^t X^t = XW_S^2 X^t \tag{3.3.6}$$

Da equação (3.3.6) nota-se que o problema é estimar X e W_s de tal forma que $XW_s^2X^t$ seja aproximadamente igual a B_s. Isto, em termos práticos, quer dizer que se tomamos um elemento de ambas igualdades da equação (3.3.6) elas podem ser expressas da seguinte maneira:

$$b_{ijs} = \sum_{k=1}^{K} x_{ik} x_{jk} w_{ks}^2 + \varepsilon_{ijs} \tag{3.3.7}$$

onde: ε_{ijs} é um erro pequeno.

Para estimar esses parâmetros uma alternativa é utilizar o método de quadrados mínimos alternados. Krazanowsk (1994) diz que a estimação do modelo é uma generalização do problema de Two-Way, tratado por Eckart e Young (1936), para procurar os valores x_{ik}, x_{jk} e w_{ks}^2, a partir do conjunto de valores conhecidos b_{ijs}.

Como mencionado, uma alternativa para a solução é utilizar o método de quadrados mínimos, da seguinte maneira:

Fixar x_{jk}e w_{ks}^2 com valores iniciais arbitrários e achar os valores de x_{ik} que minimizem a seguinte expressão:

$$V = \sum_{i=1}^{I} \sum_{j=1}^{I} \sum_{s=1}^{S} \left(b_{ijs} - \sum_{k=1}^{K} x_{ik} x_{jk} w_{ks}^2 \right)^2 \tag{3.3.8}$$

Mantendo os valores estimados de x_{ik} e pegando os valores originais de w_{sk}^2, achar valores de x_{jk} de forma que estes continuem minimizando V. Finalmente, para estes valores estimados x_{ik} e x_{jk} achar os valores de w_{sk}^2 para que V ainda seja minimizado.

Como pode-se notar, estas três minimizações formam uma sequência de métodos iterativos. A sequência continua até que exista uma mudança significativa dos valores x_{ik} ,x_{jk} e w_{sk}^2 entre duas sequências sucessivas.

3.3.4) Escolha da dimensão adequada baseada na bondade do ajuste

Quase sempre a dimensão é desconhecida, então usamos esse fato para obter soluções em algumas dimensões e escolher entre elas, baseado no princípio de Bondade do Ajuste dos dados, e sobretudo da interpretabilidade dos resultados. Davison (1983) considera que uma estratégia razoável é desenvolver uma estimação a priori da dimensão k*, e a partir daí obter soluções de (k*-3) a (k*+3) dimensões, e escolher dentre essas sete soluções. Mas, na prática, dificilmente esse critério é usado. A seguir descrevemos algumas das características de medidas de bondade de ajuste que usam os procedimentos ALSCAL e INDSCAL.

O procedimento INDSCAL é formalmente uma generalização em n-fatores da decomposição canônica para dois fatores de Eckart e Young (1936). O procedimento INDSCAL é realizado após uma conversão inicial das similaridades observadas em produtos escalares, obtendo-se alternadamente as estimativas de quadrados mínimos dos pesos W (para estimativa fixa da configuração X) e depois a configuração X supondo W dado.

Se o procedimento INDSCAL proposto por Caroll e Chang (1970) é usado, a medida de ajuste a ser usada é a correlação entre o produto escalar e o produto escalar estimado. O programa INDSCAL inicia a análise com uma matriz produto escalar que é calculada para cada indivíduo, utilizando a dupla centralização, como mencionados na equação (3.3.5) e (3.3.6).

Uma vez estimados as coordenadas estímulos $\hat{x}_{ik}$ e estimadas as ponderações para cada indivíduo $\hat{w}_{ks}$, o produto escalar estimado pode ser calculado substituindo-se a coordenada dos estímulos estimados $\hat{x}_{ik}$ e o peso estimado para cada indivíduo $\hat{w}_{ks}$ na equação (3.3.7), obtendo-se:

$$\hat{b}_{ijs} = \sum_{k=1}^{K} \hat{x}_{ik}\hat{x}_{jk}\hat{w}_{ks}^2 \qquad (3.3.8)$$

Na forma matricial, a equação (3.3.8) pode ser escrita como:

$$\hat{B}_s = \hat{X}\hat{W}_s^2\hat{X}^t \qquad (3.3.9)$$

A medida de ajuste usada pelo algoritmo INDSCAL é a **correlação simples entre o produto escalar original b_{rsi} e o produto escalar estimado** $\hat{b}_{rsi}$, isto é,

$$r = \frac{\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{S}\left(b_{ijs} - b...\right)\left(\hat{b}_{ijs} - \hat{b}...\right)}{\left[\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{S}\left(b_{ijs} - b...\right)^2 \sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{S}\left(\hat{b}_{ijs} - \hat{b}...\right)^2\right]^{\frac{1}{2}}} \qquad (3.3.10)$$

Na equação (3.3.10) $b...$ é a média total do produto escalar original, isto é,

$$b... = \frac{1}{I^2 S}\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{S} b_{ijs} \qquad (3.3.11)$$

da mesma forma $\hat{b}...$ é média total do produto escalar estimado, isto é,

$$\hat{b}... = \frac{1}{I^2 S}\sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{s=1}^{S} \hat{b}_{ijs} \qquad (3.3.12)$$

A medida de ajuste também pode ser calculada para cada indivíduo. Isto será a correlação entre o produto escalar original e o produto escalar estimado para cada indivíduo s:

$$r_s = \frac{\sum_{i=1}^{I}\sum_{j=1}^{I}\left(b_{ijs} - b..s\right)\left(\hat{b}_{ijs} - \hat{b}..s\right)}{\left[\sum_{i=1}^{I}\sum_{j=1}^{I}\left(b_{ijs} - b..s\right)^2 \sum_{r=1}^{n}\sum_{s=1}^{n}\left(\hat{b}_{ijs} - \hat{b}..s\right)^2\right]^{\frac{1}{2}}} \qquad (3.3.13)$$

Na equação (3.3.13) $b..s$ e $\hat{b}..s$ são a média do produto escalar original e estimado, respectivamente, para cada indivíduo.

Por outro lado, o procedimento ALSCAL foi desenvolvido por Takane, Young e DeLeeuw (1976). Este procedimento é o mais comum, sendo mencionado na maioria dos softwares estatísticos, e trabalha com uma medida de ajuste chamada S-STRESS (devido a sua semelhança com o STRESS) dado por:

$$S\text{-}STRESS = \left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(d_{ij}^{2} - \hat{d}_{ij}^{2} \right)^{2}}{\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^{4}} \right)^{\frac{1}{2}}$$

Nota-se que as medidas de STRESS (MDS não métrico) e S-STRESS (Modelo Euclidiano Ponderado) decrescem com a melhora dos dados ajustados, entre tanto a correlação **r** é uma medida que aumenta quando o ajuste dos dados melhora. No entanto, a medida de ajuste não é única nem suficiente para decidir a dimensão. Na literatura existem diversas pesquisas sobre este procedimento, as mais comums são os trabalhos de de Levy (1981), Takane, Young e DeLeeuw (1980) e Young, DeLeeuw e Takane (1976).

3.4.- MDS Métrico a partir de uma matriz retangular de dados de preferência ou semelhança.

Até agora vimos análise de matrizes quadradas. Na prática é comum a análise sobre matrizes retangulares. Mostraremos como aplicar Escalonamento Multidimensional (MDS) em tais matrizes baseados no trabalho apresentado por Dillon e Goldstein (1984). Muitas vezes queremos medir a preferência de determinados estímulos por indivíduos. As avaliações de cada indivíduo podem estar em quaisquer das escalas de mensuração (razão, intervalar, ordinal ou nominal). O mais comum é tê-las numa escala ordinal.

Suponhamos que temos S indivíduos e I estímulos. A cada indivíduo se apresenta os I estímulos e se pede que uma nota δ_{si} (para todo s=1,2,...S i=1,2,...,I) seja dada, de acordo com a preferência por cada estímulo. Tais preferências podem ser dispostas numa matriz, da seguinte forma:

	ESTÍMULOS				
INDIVÍDUOS		**E1**	**E2**	...	**EI**
	1	δ_{11}	δ_{12}	...	δ_{1I}
	2	δ_{21}	δ_{22}	...	δ_{2I}
	...	...	...	...	...
	S	δ_{S1}	δ_{S2}	...	δ_{SI}

A partir desta matriz de preferências podemos achar as distâncias entre indivíduos e as distâncias entre estímulos. A partir de tais matrizes de distâncias, aplicando o MDS métrico ou não métrico podemos estimar uma nova configuração de pontos tanto para os estímulos, como para os indivíduos na mesma dimensão. Isto é, seja a matriz X no espaço R^k a configuração de pontos estimados para os indivíduos. Seja a matriz Y no espaço R^k a configuração de pontos estimados para os estímulos. Neste caso, se os estímulos são considerados como variáveis aleatórias, então a solução acima poderia ser comparada com o método Q ou R na Análise Multivariada (ver apêndice B). As configurações de pontos descritas acima poderiam ser descritas matricialmente da seguinte forma:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{S1} & x_{S2} & \cdots & x_{Sk} \end{pmatrix}_{Sxk} \qquad Y = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1k} \\ y_{21} & y_{22} & \cdots & y_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ y_{I1} & y_{I2} & \cdots & y_{Ik} \end{pmatrix}_{Ixk}$$

Como as matrizes X e Y encontram-se na mesma dimensão, podemos achar a distância Euclidiana entre as linhas dessas matrizes. Isto implica que podemos achar a distância Euclidiana entre indivíduos e estímulos, da seguinte maneira:

$$d_{is} = \left[\sum_{p=1}^{k} \left(x_{ip} - y_{sp} \right)^2 \right]^{\frac{1}{2}}$$

Por outro lado se k=2, pode-se representar os pontos das matrizes X e Y em um plano de coordenadas, denotando assim o comportamento conjunto de indivíduos e estímulos em um mesmo plano. Este estudo de matrizes de preferência é conhecido como " Análise de Espaços Unidos" (joint-space analysis). Uma interpretação usual neste tipo de dados é a seguinte: suponha que d_{34} é a menor distância euclidiana ajustada dentre todas as distâncias. Isto estaria indicando a maior preferência do pelo indivíduo 3 estímulo 4.

3.5.- Generalização do Modelo Euclidiano Ponderado

Nesta seção daremos uma idéia, de forma resumida, de um modelo que é mais geral do que o modelo Euclidiano Ponderado. Os primeiros trabalhos sobre este modelo geral foram

propostos por Caroll (1972) e Tucker (1972), que o denominaram de Modelo Três Modos (The Three-mode model).

Como no modelo Euclidiano Ponderado, neste modelo também é assumido a existência de uma matriz grupo de estímulos X (isto é, esta matriz X é única para todos os indivíduos). De acordo com este modelo geral, a relação entre as matrizes X e X_s é mais complexa do que a do Modelo Euclidiano Ponderado, como descrevemos a seguir.

Do modelo Euclidiano Ponderado, equação (3.3.2), sabemos que:

$$X_S = XW_S \tag{3.5.1}$$

O modelo Três modos é expresso da seguinte maneira:

$$X_S = XW_ST_S \tag{3.5.2}$$

onde:

W_s : é uma matriz diagonal de pesos (esta matriz é igual à matriz definida no Modelo Euclidiano Ponderado).

T_s : é considerada uma matriz de transformação.

A equação (3.5.2) pode ser interpretada como as coordenadas no espaço do indivíduo, que podem ser obtido através da combinação, ponderação e rotação das coordenadas no espaço grupo. Nota-se que o modelo Euclidiano Ponderado é um caso particular do modelo Três modos, pois basta fazer T_s=I. Em termos práticos, a matriz de transformação T_s basicamente produz uma rotação nos eixos originais.

O modelo em termos de distância pode ser considerado como:

$$d_{ijs} = \left[\sum_{k=1}^{K} \left(x_{iks} - x_{jks} \right)^2 \right]^{\frac{1}{2}} \tag{3.5.3}$$

onde:

x_{iks} : é a coordenada do estímulo **i** na dimensão **k** do indivíduo **s**.

x_{jks} : é a coordenada do estímulo **j** na dimensão **k** do indivíduo **s**.

Se considerarmos a matriz $R_s=T_sT_s^t$, com elementos $r_{kk's}$, então a equação (3.5.3) pode ser escrita em função dos elementos da matriz (X_s), da seguinte maneira:

$$\left(\delta_{ijs}\right)^{\frac{1}{2}} = \sum_{k=1}^{K} w_{ks}^2\left(x_{ik}-x_{jk}\right)^2 + \sum_{k=1}^{K}\sum_{k'=1(k'\neq k)}^{K} w_{ks}w_{k's}r_{kk's}\left(x_{ik}-x_{jk}\right)\left(x_{ik'}-x_{jk'}\right) \tag{3.5.4}$$

Se $r_{kk's}=0$, então o modelo (3.5.4) reduz-se ao modelo Euclidiano Ponderado. A somatória dupla na equação (3.5.4) sobre as dimensões k e k' representa a soma das interações entre as diferenças das coordenadas sobre todos os possíveis pares de dimensões. Os parâmetros $r_{kk's}$ caracterizam o tamanho e a direção da interação entre as diferenças das coordenadas das dimensões k e k' nos julgamentos do indivíduo s, pois o sinal de $r_{kk's}$ indica a direção da interação e o valor absoluto indica o tamanho daquela interação.

Tuckers (1972) afirma que a matriz de pesos W_s e a matriz de transformação T_s não representam a mais concisa descrição das características individuais. A parte de estimação não será desenvolvida, pois, na prática, os pesquisadores optam pelo Modelo Euclidiano Ponderado. Por outro lado, Caroll (1972) baseia-se no modelo Três Modos para formular o Modelo de preferência Ponto Ideal Geral, que será descrito detalhadamente no capítulo IV.

Capítulo IV

MODELOS DE PREFERÊNCIA

4.1.- Introdução

Neste capítulo, estamos interessados em apresentar os modelos de preferência baseados na teoria proposta por Caroll (1972), Davison (1983), Artes (1991) e Warren (1992). Os modelos de preferência surgem com a idéia do modelo de distância para dados de preferência muitas vezes conhecidos como Modelos Desdobrados (Unfolding Models), assim chamados por Coombs (1964). Caroll (1972) considera este tipo de análise como uma Análise Externa de MDS, porque se inicia com dados de preferência para cada indivíduo e com a suposição de que já conhecemos uma configuração de pontos (que neste trabalho são as variáveis não observáveis da solução do MDS). A partir daí formula-se modelos para cada indivíduo, os quais podem ser ajustados usando técnicas da Análise de Regressão Linear Múltipla ou Regressão Monotônica ou Isotônica, e através desse procedimento encontrar o ponto de preferência máxima de cada indivíduo. Por outro lado, Davison (1983) considera a análise como Modelo de distância para Preferências, mas baseado na idéia de Caroll. Observa-se que os modelos de preferência são exclusivamente para dados de preferência individuais.

Neste capítulo apresentamos os quatro modelos de preferência. O primeiro modelo é o de Vetor de Preferência, considerado modelo linear, e os seguintes são os modelos Ponto Ideal ou Euclidiano Simples, Ponderado e Geral, considerados modelos não lineares. Na segunda parte, faremos a estimação métrica e não métrica dos modelos. Neste capítulo trabalharemos basicamente sobre a solução que é obtida através da técnica de MDS (métrico, não métrico, ou euclidiano ponderado).

4.2.- Problema Geral dos Modelos de Preferência

Partamos da suposição de que temos I estímulos, S indivíduos e δ_{is} indicando a preferência do indivíduo s pelo estímulo i (s=1,2,...,S e i=1,2,...,I). A partir desta informação, podemos encontrar uma configuração de pontos $\hat{X}$ de ordem (IxK) para poder expressar o comportamento dos estímulos utilizando a técnica do MDS. Para efeitos de notação, a matriz $\hat{X}$ será denotada simplesmente por X. Esta matriz X é muito importante, é a partir dela que serão formulados os modelos de preferência. Fazendo analogia com o trabalho de Artes (1991), a matriz X é considerada como uma matriz de planejamento, e com respeito à Analise de Regressão ela é considerada como a matriz das regressoras (variáveis independentes) em relação às respostas individuais de preferências δ_{is}, que são consideradas como variável resposta. Então vamos definir a matriz X da seguinte maneira:

$$X_{IxK} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{I1} & x_{I2} & \cdots & x_{Ik} \end{pmatrix} \begin{matrix} \text{estímulo.1} \\ \text{estímulo.2} \\ \vdots \\ \text{estímulo.I} \end{matrix}$$

Pode-se notar que a matriz X descreve o comportamento dos I estímulos em K dimensões para os S indivíduos. Na prática, essas dimensões podem ser consideradas como atributos, fatores ou variáveis não observáveis, pois, por definição, os atributos são considerados combinações dos estímulos. O único problema, que sempre acontece na prática, é tratar de identificar essas dimensões mediante nomes específicos que serão tratados como atributos. Para o desenvolvimento dos Modelos de Preferência já se pressupõe conhecida a matriz X, e sobre essa matriz formularemos modelos para poder encontrar o ponto de preferência máxima para cada indivíduo.

Na prática, também recomenda-se encontrar uma configuração de pontos para os indivíduos, isto é, achar uma matriz de distância para os indivíduos e a partir daí encontrar a configuração de pontos que explique graficamente o comportamento dos mesmos, para depois ter uma idéia da relação existente entre estes indivíduos e os estímulos (ver seção 3.4). Por outro lado, os Modelos de Preferência podem ser formulados nas escalas de mensuração intervalar, razão, nominal ou ordinal. Diante do exposto, estudaremos os Modelos de Preferência.

4.3.- Definição dos Modelos de Preferência

4.3.1.- Modelo de Preferência Vetor.

O modelo de preferência Vetor foi analisado por Caroll (1972) e Bechtel, Tucker e Chang (1971). A única suposição forte é considerar que as preferências (δ_{is}) estão relacionadas linearmente com as dimensões (ou atributos). Por esta razão, o modelo de Preferência Vetor é considerado modelo linear. A formulação estatística foi proposta da seguinte maneira:

$$\delta_{is} = \sum_{k=1}^{K} x_{ik}\beta_{ks} + \varepsilon_s \qquad (4.3.1)$$

onde:

δ_{is} : é a preferência do indivíduo **s** pelo estímulo **i**.
x_{ik} : é a coordenada para o estímulo **i** no atributo (ou dimensão) **k.**
β_{ks} : é o parâmetro a se estimar do indivíduo **s** na dimensão **k**.
ε_s : é uma constante aditiva única para o indivíduo **s**.
Para todo: i=1,2,...,**I** , **s**=1,2,...,**S** e k=1,2,...,**K**

A equação (4.3.1) pode ser escrita na forma matricial da seguinte maneira:

$$\delta_S = X\beta_S + e_S \qquad (4.3.2)$$

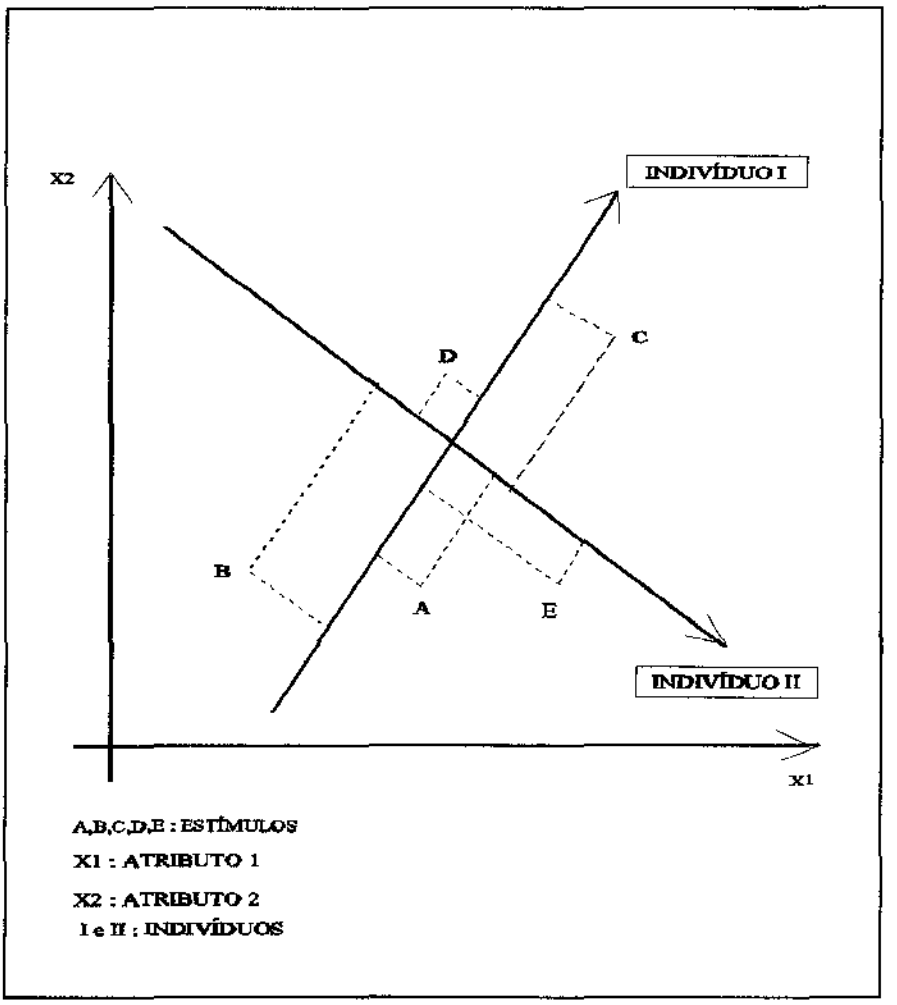
onde:

$$\delta_S = \begin{pmatrix} \delta_{1s} \\ \vdots \\ \delta_{is} \\ \vdots \\ \delta_{Is} \end{pmatrix}_{Ix1} \qquad X = \begin{pmatrix} x_{11} & \cdots & x_{1k} & \cdots & x_{1K} \\ \vdots & \ddots & \vdots & & \vdots \\ x_{i1} & & x_{ik} & & x_{iK} \\ \vdots & & \vdots & \ddots & \vdots \\ x_{I1} & \cdots & x_{Ik} & \cdots & x_{IK} \end{pmatrix}_{IxK} \qquad \beta_S = \begin{pmatrix} \beta_{1s} \\ \vdots \\ \beta_{ks} \\ \vdots \\ \beta_{Ks} \end{pmatrix}_{Kx1}$$

A equação (4.3.2) é a mesma usada numa Regressão Linear Múltipla, exceto que o erro neste caso é único, mas ainda, é considerada uma constante. Também percebe-se que temos uma equação de regressão para cada indivíduo. As estimativas dos parâmetros β_s podem ser feitas através de quadrados mínimos, e a interpretação é a mesma que em uma regressão linear múltipla, isto é, os parâmetros β_s representam a importância de cada atributo na preferência do estímulo pelo indivíduo.

Tudo é baseado na seguinte idéia de Caroll (1972): Suponhamos que temos 5 estímulos (A,B,C,D,E), 2 dimensões (ou atributos) x_1 ,x_2 e 2 indivíduos, isto é: I=5, K=2 e S=2. Consideremos o espaço dos estímulos (o seguinte gráfico pode ser feito sobre os eixos dos atributos):

Gráfico 4.1. Esquema do Modelo de Preferência Vetor.



A idéia básica dessa interpretação é imaginar nesse espaço a existência de vetores, cada um associado a um indivíduo, que indicam a direção do crescimento de sua preferência. Dessa forma, a ordem de preferência dos estímulos para um determinado sujeito seria dada pela projeção dos estímulos em seu correspondente vetor. Por exemplo, do gráfico 4.1 é possível observar que o indivíduo 1 prefere o estímulo C, seguido pelos estímulos D, E, A e B. Analogamente a ordem de preferências (de maior a menor preferência) dos estímulos para o indivíduo 2 é E, C, A, D e B.

4.3.2.- Modelos de Preferências Ponto Ideal ou Euclidiano

Os modelos Ponto Ideal podem ser expressos de três formas diferentes, em ordem hierárquica como Simples, Ponderado e Geral. Estes modelos não precisam da suposição de linearidade, daí que são considerados modelos não lineares. Eles partem do princípio de que para cada indivíduo existe um estímulo, que é o de preferência máxima (ponto ideal), e que, quanto mais parecido com ele for o produto, maior será sua preferência.

4.3.2.1.- Modelo Ponto Ideal Simples (P.I.S).

Considerado o modelo mais elementar, parte do princípio de que todas as dimensões (ou atributos) são igualmente importantes. A única mudança que pode ocorrer é no peso estimado para cada indivíduo. Este é o primeiro modelo que assume função não linear e é apresentado da seguinte maneira:

$$\delta_{is} = \sum_{k=1}^{K} w_s\left(x_{ik} - \pi_{sk}\right)^2 + \varepsilon_s^* \qquad (4.3.3)$$

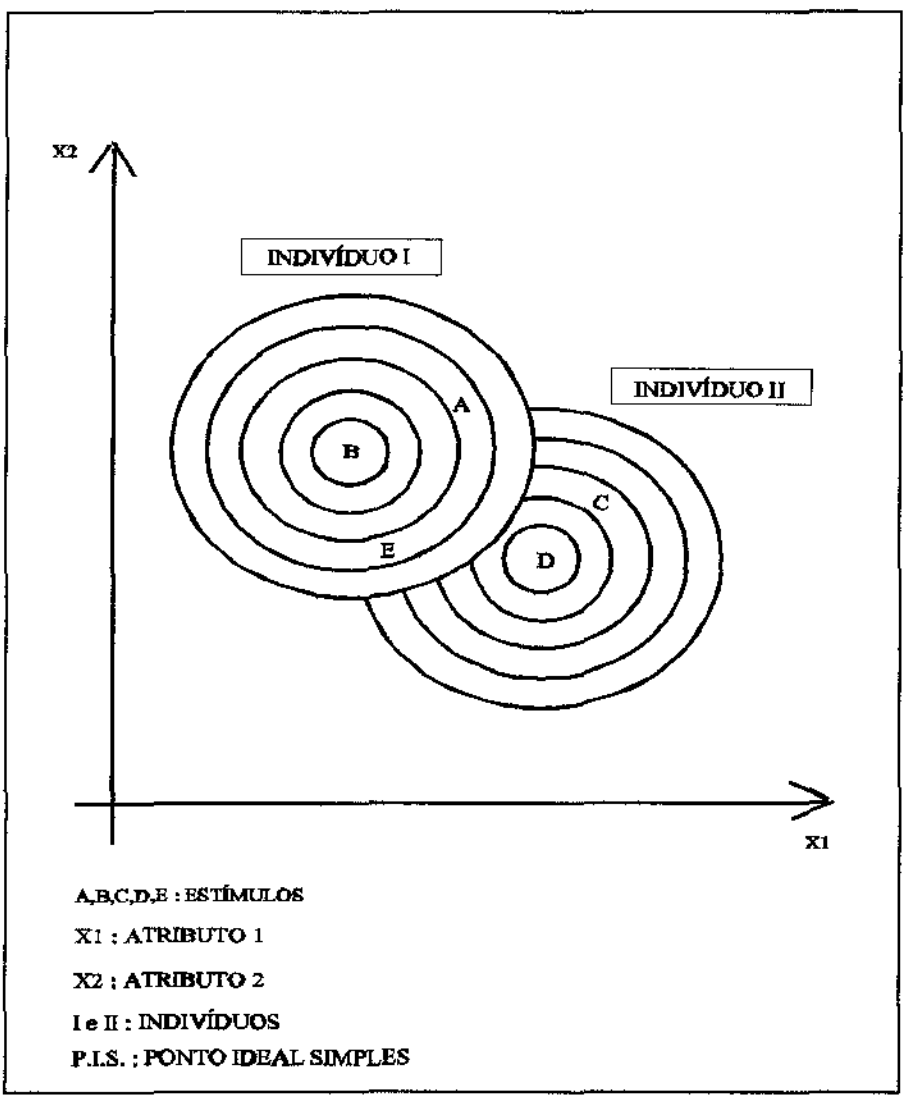
onde:

δ_{is} : é a preferência do indivíduo **s** pelo estímulo **i**.
x_{ik} : é a coordenada para o estímulo **i** no atributo (ou dimensão) **k**.
π_{sk} : é o ponto ideal a estimar do indivíduo **s** na dimensão **k**.
w_s : é o peso a ser estimado para o indivíduo **s**.
ε_s^* : é uma constante aditiva única para o indivíduo **s**.
Para todo: i=1,2,...,**I** , s=1,2,...,**S** e k=1,2,...,**K**

A idéia deste modelo é formar curvas de isopreferência. No gráfico 4.2 as circunferências são interpretadas como linhas de isopreferência e o centro são os pontos ideais.

Retomando o exemplo do Modelo de preferência vetor temos 5 estímulos (A,B,C,D,E), 2 dimensões (ou atributos) x_1 ,x_2 e 2 indivíduos, isto é: I=5, K=2 e S=2. Consideremos o espaço dos estímulos representado no gráfico a seguir, que pode ser feito sobre os eixos dos atributos:

Gráfico 4.2. Linhas de Isopreferência de um Modelo Ponto Ideal Simples.



Nota-se nesse espaço a existência de circunferências, as quais denominaremos de linhas de isopreferência, associadas a cada indivíduo que indicam uma ordem na preferência pelos

estímulos. Do gráfico 4.2 temos que, para o indivíduo I os estímulos mais preferidos encontram-se entre B, A e E, e os estímulos com menor preferência são D e C. Analogamente, para o indivíduo II temos que os estímulos de maior preferência são D, C e E, e os estímulos menos preferidos são A e B.

Na parte de estimação do modelo, mostraremos que o modelo de Preferência Vetor é considerado um caso particular do modelo Ponto Ideal Simples.

4.3.2.2.- Modelo Ponto Ideal Ponderado (P.I.P).

Os modelos ponto ideal simples e ponto ideal ponderado diferem apenas na forma de definição dos pesos w_{ks}, isto é, vimos que o Modelo Ponto Ideal Simples assume igual ponderação para todos os atributos (ou dimensões), enquanto que o Modelo Ponto Ideal Ponderado permite uma ponderação diferente para cada dimensão (ou atributo). O modelo é apresentado da seguinte maneira:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks}\left(x_{ik} - \pi_{sk}\right)^2 + \varepsilon_s^* \qquad (4.3.4)$$

onde:

δ_{is} : é a preferência do indivíduo **s** pelo estímulo **i**.
x_{ik} : é a coordenada para o estímulo **i** no atributo (ou dimensão) **k**.
π_{sk} : é o ponto ideal a estimar do indivíduo **s** na dimensão **k**.
w_{ks} : é o peso a estimar para o indivíduo **s** na dimensão **k**.
ε_s^* : é uma constante aditiva única para o indivíduo **s**.

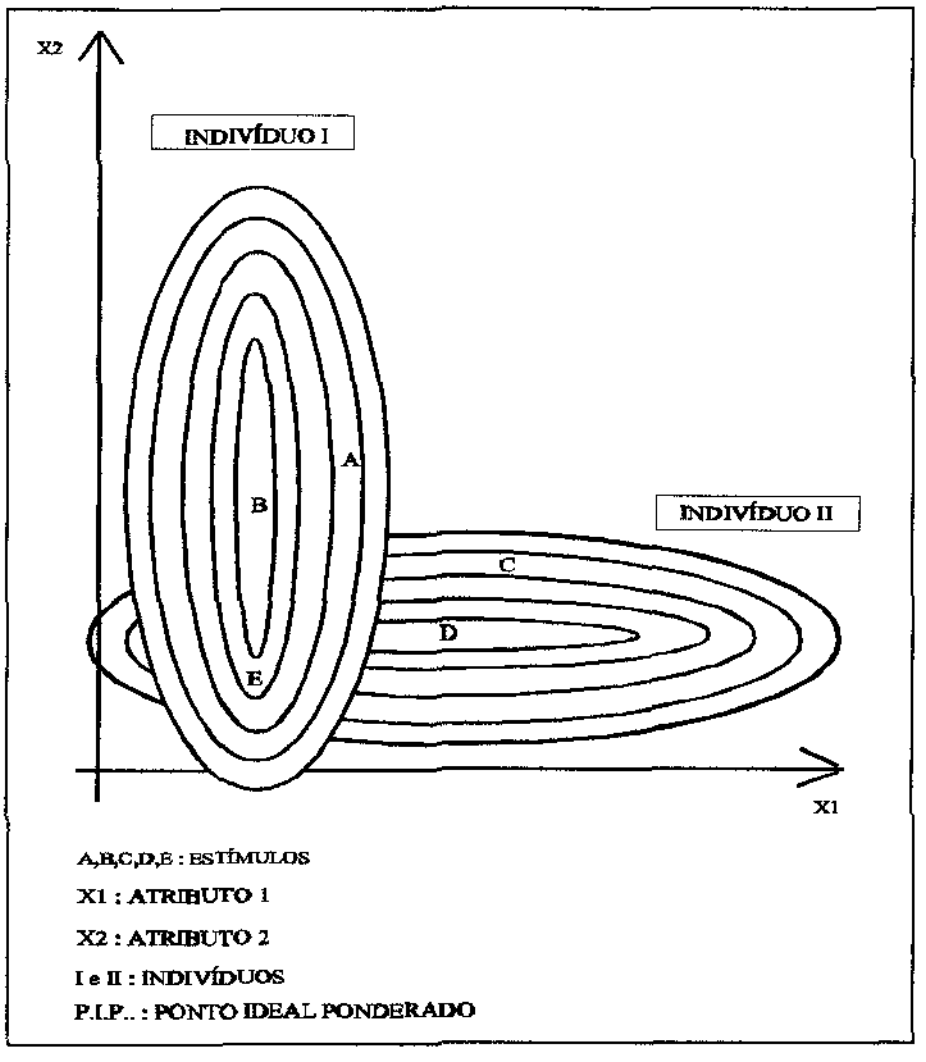
Para todo: i=1,2,...,**I** , **s**=1,2,...,**S** e k=1,2,...,**K**

Pode-se notar que a equação (4.3.4) é similar ao modelo Euclidiano Ponderado para similaridades (ver equação 3.3.4). As únicas diferenças são o ponto ideal π_{sk} e a constante aditiva ε_s^*.

A seguir mostramos a idéia gráfica do Modelo Ponto Ideal Ponderado. No gráfico 4.3 observa-se como ficam as linhas de isopreferência para 2 indivíduos com 2 atributos. Agora

estas linhas são elipses, cujo centro continua sendo o ponto ideal. Retomando o exemplo anterior, temos:

Gráfico 4.3. Linhas de isopreferência de um Modelo Ponto Ideal Ponderado



A idéia basica do modelo é imaginar nesse espaço a existência de elipses associadas a cada indivíduo, os quais indicam a ordem da preferência pelos estímulos, assim o gráfico mostra que o indivíduo I prefere os estímulos B, E e A, e com menor preferência os estímulos D e C. Analogamente, para o indivíduo II, temos que os estímulos mais preferidos são D, C e E, e os estímulos com menor preferência são A e B.

Na parte de estimação do modelo mostraremos que o modelo de Preferência Vetor e o modelo Ponto Ideal Simples são casos particulares do modelo Ponto Ideal Ponderado.

4.3.2.3.- Modelo Ponto Ideal Geral (P.I.G).

O modelo Ponto Ideal Geral é considerado como uma extensão do modelo para similaridades proposto por Tuckers (1972), (ver capítulo III, seção 3.6). A seguir o modelo é apresentado baseado na forma modificada dada por Davison (1983).

Este modelo assume que existe uma matriz (KxK) de transformação T_s e uma matriz diagonal de pesos W_s para cada indivíduo. A matriz T_s é considerada uma matriz de transformação linear, a qual produz uma rotação nos eixos originais. A interação é traduzida como uma mudança da direção nos eixos da elipse (ver figura 4.4). Aplicando estas matrizes sobre a matriz **X**, é produzida uma nova matriz de coordenadas para cada indivíduo, isto é:

$$X_S = XW_ST_S \tag{4.3.5}$$

Descrevendo as matrizes da equação (4.3.5) temos:

$$(X_s)_{IxK} = \begin{pmatrix} x_{11s} & \cdots & x_{1ks} & \cdots & x_{1Ks} \\ \vdots & \ddots & \vdots & & \vdots \\ x_{i1s} & \cdots & x_{iks} & \cdots & x_{iKs} \\ \vdots & & \vdots & \ddots & \vdots \\ x_{I1s} & \cdots & x_{Iks} & \cdots & x_{IKs} \end{pmatrix} \begin{matrix} \vec{x}^{*t}_1 \\ \vdots \\ \vec{x}^{*t}_i \\ \vdots \\ \vec{x}^{*t}_I \end{matrix}$$

$$(X)_{IxK} = \begin{pmatrix} x_{11} & \cdots & x_{1k} & \cdots & x_{1K} \\ \vdots & \ddots & \vdots & & \vdots \\ x_{i1} & \cdots & x_{ik} & \cdots & x_{iK} \\ \vdots & & \vdots & \ddots & \vdots \\ x_{I1} & \cdots & x_{Ik} & \cdots & x_{IK} \end{pmatrix} \begin{matrix} \vec{x}_1 \\ \vdots \\ \vec{x}_i \\ \vdots \\ \vec{x}_I \end{matrix}$$

$$(W_s)^*_{KxK} = \begin{pmatrix} w^*_{1s} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & w^*_{ks} & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & w^*_{Ks} \end{pmatrix}$$

$$\left(T_s^*\right)_{KxK} = \begin{pmatrix} t_{11s}^* & t_{12s}^* & \cdots & t_{1Ks}^* \\ t_{21s}^* & t_{22s}^* & \cdots & t_{2Ks}^* \\ \vdots & \vdots & \cdots & \vdots \\ t_{k1s}^* & t_{k2s}^* & \cdots & t_{KKs}^* \end{pmatrix}$$

O modelo de preferências (4.3.5) pode também ser escrito da seguinte forma:

$$\delta_{is} = \sum_{k=1}^{K} \left(x_{iks} - \pi_{sks}\right)^2 + \varepsilon_s^* \qquad (4.3.6)$$

A equação (4.3.6) está indicando a distância entre o estímulo **i** e o ponto ideal do indivíduo **s**. Consequentemente a equação (4.3.6) expressa o modelo Ponto Ideal Geral em termos da resposta de cada indivíduo e as coordenadas para os estímulos dada pela matriz $\mathbf{X_s}$.

onde:

δ_{is} : é a preferência do indivíduo **s** pelo estímulo **i**.

x_{iks} : é a coordenada para o estímulo **i** no atributo (ou dimensão) **k** em $\mathbf{X_s}$.

π_{sks} : é o ponto ideal a ser estimado do indivíduo **s** na dimensão **k** em $\mathbf{X_s}$.

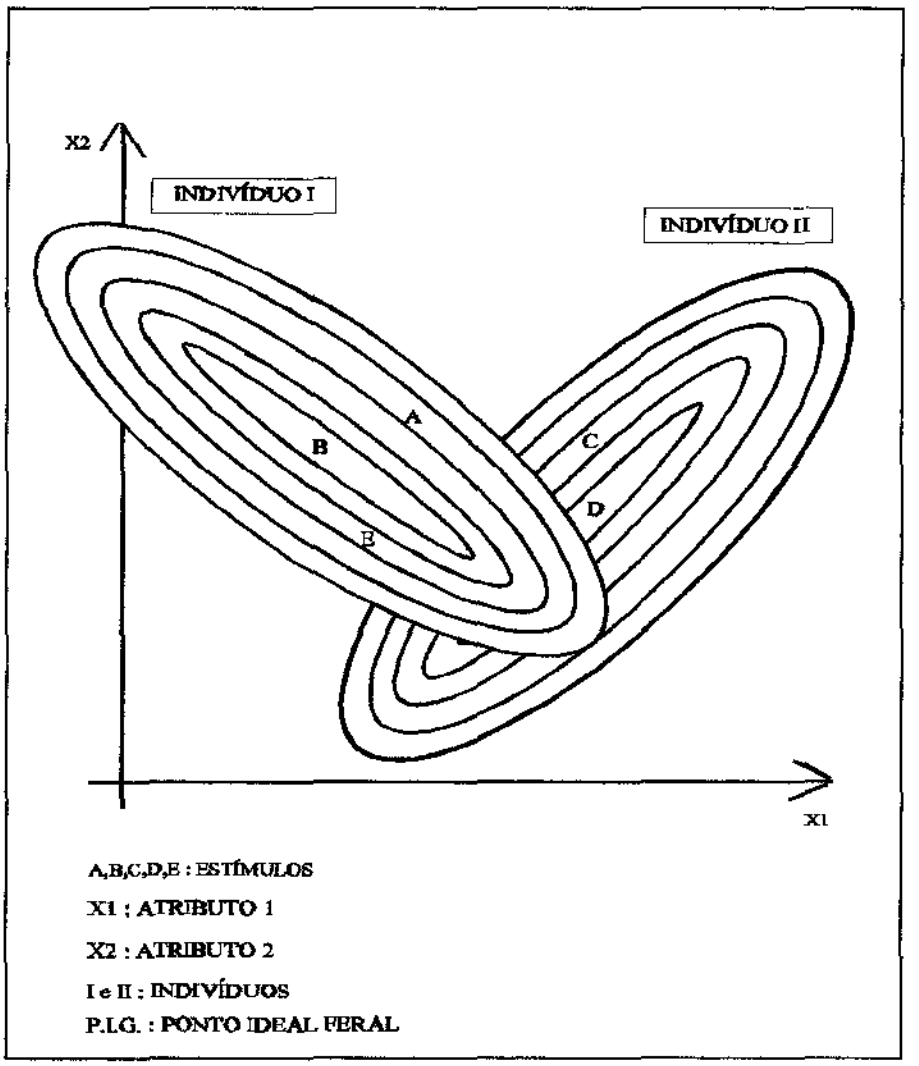
ε_s^* : é uma constante aditiva única para o indivíduo **s**.

Para todo: i=1,2,...,**I** , s=1,2,...,**S** e k=1,2,...,**K**

No gráfico 4.4 mostramos a idéia do Modelo Ponto Ideal Geral, onde percebemos a mudança nos eixos das elipses associadas a cada indivíduo.

BIBLIOTECA CENTRAL

Gráfico 4.4. Linhas de Isopreferência de um Modelo Ponto Ideal Geral.



Como visto anteriormente, neste modelo a idéia é imaginar a interação entre as dimensões (ou atributos) como uma rotação nos eixos, onde através das elipses pode-se perceber a ordem das preferências pelos estímulos associadas a cada indivíduo. Do gráfico 4.4 vemos que o indivíduo I tem maior preferência pelos estímulos B, E e A, e os menos preferidos são D e C. Para o indivíduo II os estímulos mais preferidos são D, C e E e os estímulos menos preferidos são A e B.

A seguir descrevemos o procedimento para simplificar o Modelo Ponto Ideal Geral. A equação (4.3.6) pode ser escrita vetorialmente como:

$$\delta_{is} = \left(\vec{x}_i^* - \vec{\pi}_s^*\right)^t \left(\vec{x}_i^* - \vec{\pi}_s^*\right) + \varepsilon_s^* \tag{4.3.7}$$

Pode-se perceber que a equação (4.3.7) está escrita em termos da matriz $\mathbf{X_s}$. Agora vamos a escrever tal equação em termos da matriz X.

Da equação (4.3.5), multiplicando as matrizes temos:

$$\left(X_s\right)_{IxK} = \begin{pmatrix} \sum_{k=1}^{K} x_{1k} w_{ks}^* t_{k1s}^* & \sum_{k=1}^{K} x_{ik} w_{ks}^* t_{k2s}^* & \cdots & \sum_{k=1}^{K} x_{1k} w_{ks}^* t_{kks}^* \\ \sum_{k=1}^{K} x_{2k} w_{ks}^* t_{k1s}^* & \sum_{k=1}^{K} x_{2k} w_{ks}^* t_{k2s}^* & \cdots & \sum_{k=1}^{K} x_{2k} w_{ks}^* t_{kks}^* \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{K} x_{Ik} w_{ks}^* t_{k1s}^* & \sum_{k=1}^{K} x_{Ik} w_{ks}^* t_{k2s}^* & \cdots & \sum_{k=1}^{K} x_{Ik} w_{ks}^* t_{kks}^* \end{pmatrix}$$

e escrevendo vetorialmente:

$$\left(X_s\right)_{IxK} = \begin{pmatrix} \vec{x}_1^{*t} \\ \vec{x}_2^{*t} \\ \vdots \\ \vdots \\ \vec{x}_I^{*t} \end{pmatrix} = \begin{pmatrix} \left[\left(T_s^*\right)\left(W_s^*\right)\vec{x}_1\right]^t \\ \left[\left(T_s^*\right)\left(W_s^*\right)\vec{x}_2\right]^t \\ \vdots \\ \vdots \\ \left[\left(T_s^*\right)\left(W_s^*\right)\vec{x}_I\right]^t \end{pmatrix} \tag{4.3.8}$$

Da equação (4.3.8) vemos que a relação entre a matriz **X** e a matriz $\mathbf{X_s}$ é:

$$\vec{x}_i^* = \left(T_s^*\right)\left(W_s^*\right)\vec{x}_i \tag{4.3.9}$$

analogamente, pode-se mostrar que:

$$\vec{\pi}_S^* = \left(T_S^*\right)\left(W_S^*\right)\vec{\pi}_S \tag{4.3.10}$$

Substituindo as equações (4.3.9) e (4.3.10) na equação (4.3.7) temos:

$$\delta_{is} = \left[\left(T_S^*\right)\left(W_S^*\right)\vec{x}_i - \left(T_S^*\right)\left(W_S^*\right)\vec{\pi}_S\right]^t\left[\left(T_S^*\right)\left(W_S^*\right)\vec{x}_i - \left(T_S^*\right)\left(W_S^*\right)\vec{\pi}_S\right] + e_S^* \tag{4.3.11}$$

Operando na equação (4.3.11) temos:

$$\delta_{is} = \left(\vec{x}_i - \vec{\pi}_S\right)^t\left(W_S^*\right)\left(T_S^{*t}\right)\left(T_S^*\right)\left(W_S^*\right)\left(\vec{x}_i - \vec{\pi}_S\right) + \varepsilon_S^* \tag{4.3.12}$$

Fixando:

$$R_S^* = T_S^{*t} T_S^* \tag{4.3.13}$$

e substituindo na equação (4.3.12), temos:

$$\delta_{is} = \left(\vec{x}_i - \vec{\pi}_S\right)^t\left(W_S^*\right)\left(R_S^*\right)\left(W_S^*\right)\left(\vec{x}_i - \vec{\pi}_S\right) + \varepsilon_S^* \tag{4.3.14}$$

Multiplicando as matrizes e vetores da equação (4.3.14), obtemos o seguinte:

$$\begin{aligned}
\delta_{is} = & \sum_{k=1}^{K}\left(x_{ik} - \pi_{sk}\right)^2 w_{ks}^* r_{kks}^* w_{ks}^* + \\
& + \sum_{k=2(k\neq 1)}^{K}\left(x_{i1} - \pi_{s1}\right)\left(x_{ik} - \pi_{sk}\right) w_{ks}^* r_{k1s}^* w_{1s}^* + \\
& + \sum_{k=1(k\neq 2)}^{K}\left(x_{i2} - \pi_{s2}\right)\left(x_{ik} - \pi_{sk}\right) w_{ks}^* r_{k2s}^* w_{2s}^* + \\
& + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + \\
& + \sum_{k=1(k\neq K)}^{K-1}\left(x_{iK} - \pi_{sK}\right)\left(x_{ik} - \pi_{sk}\right) w_{ks}^* r_{kKs}^* w_{Ks}^* + e_s^*
\end{aligned} \tag{4.3.15}$$

e mais ainda, a equação (4.3.15) pode ser escrita da seguinte maneira:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks}^{2*} r_{kks}^{*} \left(x_{ik} - \pi_{sk}\right)^2 + \\ + \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{*} w_{k's}^{*} r_{kk's}^{*} \left(x_{ik} - \pi_{sk}\right)\left(x_{ik'} - \pi_{sk'}\right) + e_{s}^{*} \tag{4.3.16}$$

isto mostra que a equação (4.3.6) pode ser escrita como a equação (4.3.16).

Seja a seguinte mudança de variável:

$$w_{ks} = w_{ks}^{2*} r_{kks}^{*} \tag{4.3.17}$$

Da equação (4.3.17) pode-se mostrar que:

$$w_{ks}^{*} = \frac{w_{ks}^{\frac{1}{2}}}{\sqrt{r_{kks}^{*}}} \tag{4.3.18}$$

$$w_{k's}^{*} = \frac{w_{k's}^{\frac{1}{2}}}{\sqrt{r_{k'k's}^{*}}} \tag{4.3.19}$$

e seja também:

$$r_{kk's} = \frac{r_{kk's}^{*}}{\sqrt{r_{kks}^{*} r_{k'k's}^{*}}} \tag{4.3.20}$$

Substituindo as equações (4.3.17), (4.3.18), (4.3.19) e (4.3.20) na equação (4.3.16), temos:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks} r_{kks} \left(x_{ik} - \pi_{sk}\right)^2 + \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \left(x_{ik} - \pi_{sk}\right)\left(x_{ik'} - \pi_{sk'}\right) + e_s^* \quad (4.3.21)$$

Então mostramos que a equação (4.3.6) pode ser escrita em termos da matriz original **X.**

onde:

δ_{is}: é a preferência do indivíduo **s** pelo estímulo **i**.

x_{ik} : é a coordenada para o estímulo **i** no atributo(ou dimensão) **k.**

π_{sk} : é o ponto ideal a ser estimado do indivíduo **s** na dimensão **k**.

$\pi_{sk'}$: é o ponto ideal a ser estimado do indivíduo **s** na dimensão **k'**.

w_{ks} : é o peso a ser estimado para o indivíduo **s** na dimensão **k**.

$w_{k's}$: é o peso a ser estimado para o indivíduo **s** na dimensão **k'**.

$r_{kk's}$: é a interação entre as dimensões (ou atributos) **k** e **k'** para o indivíduo **s**.

ε_s^* : é uma constante aditiva única para o indivíduo **s.**

Para todo: $k \neq k'$ onde: $k,k' = 1,2,\ldots\ldots\ldots\ldots,K$.

O sinal e valor absoluto de $r_{kk's}$ na equação (4.3.7) caracterizam a direção e a magnitude da interação entre as dimensões (ou atributos) **k** e **k'** nos julgamentos do indivíduo **s**, respectivamente. Também percebe-se que o modelo de Preferência Vetor, o modelo Ponto Ideal Simples e o modelo Ponto Ideal Ponderado são casos particulares do modelo Ponto Ideal Geral, pois se $r_{kk's}=0$, para todos os pares de dimensões (ou atributos) **k** e **k'** então a equação (4.3.7) é igual ao modelo Euclidiano Ponderado (ver equação 4.3.4). E além disso, se as ponderações são iguais para as K dimensões (ou atributos), isto é $w_{ks} = w_s$, então o modelo Ponto Ideal Simples é um caso particular do modelo Ponto Ideal Geral. Assim, o modelo de Preferência Vetor também é um caso particular deste.

4.4.- Estimação dos Modelos de Preferência

Nesta seção descreveremos dois métodos para estimar os modelos de preferência que estão ligados à avaliação dos indivíduos (δ_{is}). Quando a avaliação dos indivíduos (δ_{is}) estiver em uma escala razão ou intervalar, será denominada de estimação métrica. Se a avalição estiver em uma escala ordinal ou nominal, será denominada de estimação não métrica.

4.4.1. Estimação Métrica

A estimação dos modelos de preferência na parte métrica será realizada usando, basicamente, a Regressão Linear Múltipla. Para isto será analisada a parte matricial, com o propósito de construir alguns testes de hipóteses que servirão para medir a qualidade do ajuste, assim como para interpretar melhor os modelos. Assume-se que são satisfeitas todas as suposições para uma Análise de Regressão Múltipla, ver Draper e Smith (1981).

4.4.1.1.- Estimação do Modelo de Preferência Vetor

Já vimos que o modelo de Preferência Vetor foi expresso da seguinte maneira:

$$\delta_{is} = \sum_{k=1}^{K} x_{ik}\beta_{ks} + \varepsilon_s \qquad (4.4.1)$$

e na forma matricial pode ser escrito como:

$$\delta_S = X\beta_S + \varepsilon_S \qquad (4.4.2)$$

A equação (4.4.2) pode ser estimada através do método de quadrados mínimos em uma Regressão Linear Múltipla, pois estamos supondo que as K dimensões (ou atributos) estão relacionadas linearmente com as preferências (δ_s). A análise produz estimativas dos parâmetros para cada indivíduo, isto é, $\hat{\beta}_S$ e, também, $\hat{\varepsilon}_S$. Além disso, estima-se um coeficiente de correlação múltipla R_{vs}^2.

Interpretação dos parâmetros

A interpretação dos parâmetros estimados $\hat{\beta}_S$ é semelhante à de uma análise de Regressão. O coeficiente de correlação múltipla ($\mathbf{R}_{vs}^2$) indica a proporção da variância da preferência dos indivíduos que pode ser avaliada pelas k dimensões (ou atributos) no modelo e, também pode ser considerada como uma medida de ajuste para o modelo.

Por outro lado, pode-se testar a significância do coeficiente de correlação múltipla ($\mathbf{R}_{vs}^2$) mediante as seguintes hipóteses:

$$H_0 : R_{vs}^2 = 0$$
$$H_1 : R_{vs}^2 > 0$$

A estatística F a ser utilizada é:

$$F = \frac{R_{VS}^2}{\left(1 - R_{VS}^2\right)} x \frac{(I-K-1)}{K} \approx F_{[K],[I-K-1]}$$

4.4.1.2.- Estimação dos modelos Ponto Ideal ou Euclidiano

A estimação destes modelos consiste basicamente na linearização e reparametrização dos modelos originais.

- <u>Estimação do Modelo Ponto Ideal Simples (P.I.S)</u>.

Vimos que o modelo Ponto Ideal Simples é expresso da seguinte maneira (ver equação 4.3.3):

$$\delta_{is} = \sum_{k=1}^{K} w_s \left(x_{ik} - \pi_{sk}\right)^2 + \varepsilon_s^* \qquad (4.4.3)$$

Linearizando o modelo temos:

$$\delta_{is} = w_s \sum_{k=1}^{K} x_{ik}^2 + \sum_{k=1}^{K} \left(-2 w_s \pi_{sk} x_{ik}\right) + \sum_{k=1}^{K} w_s \pi_{sk}^2 + \varepsilon_s^* \qquad (4.4.4)$$

Reparametrizando, temos:

$$\varepsilon_s = \sum_{k=1}^{K} w_s \pi_{sk}^2 + \varepsilon_s^* \tag{4.4.5}$$

$$\beta_{ks} = -2 w_s \pi_{sk} \tag{4.4.6}$$

e, substituindo (4.4.5) e (4.4.6) na equação (4.4.4) temos:

$$\delta_{is} = w_s \sum_{k=1}^{K} x_{ik}^2 + \sum_{k=1}^{K} \beta_{ks} x_{ik} + \varepsilon_s \tag{4.4.7}$$

Na equação (4.4.7), se $w_s = 0$, então obtém-se o modelo de Preferência Vetor, por conseguinte vemos que o modelo de Preferência Vetor, é um caso particular do modelo ponto Ideal Simples.

Em forma matricial a equação (4.4.7) pode ser escrita da seguinte maneira:

$$\delta_s = X^e \beta_s^e + \varepsilon_s \tag{4.4.8}$$

onde:

$$\delta_s = \left(\delta_{1s} \quad \delta_{2s} \quad \cdots \quad \delta_{is} \quad \cdots \quad \delta_{Is}\right)^t_{Ix1}$$

$$X^e = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} & \cdots & x_{1K} & \sum_{k=1}^{K} x_{1k}^2 \\ x_{21} & x_{22} & \cdots & x_{2k} & \cdots & x_{2K} & \sum_{k=1}^{K} x_{2k}^2 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ik} & \cdots & x_{iK} & \sum_{i=1}^{K} x_{ik}^2 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & \vdots \\ x_{I1} & x_{I2} & \cdots & x_{Ik} & \cdots & x_{IK} & \sum_{k=1}^{K} x_{Ik}^2 \end{pmatrix}_{Ix(K+1)}$$

$$\beta_s^e = \left(\beta_{1s} \quad \beta_{2s} \quad \cdots \quad \beta_{ks} \quad \cdots \quad \beta_{Ks} \quad w_s\right)^t_{(K+1)x1}$$

A equação (4.4.8) pode ser interpretada como um modelo de Regressão Linear Múltipla e os parâmetros podem ser estimados via quadrados mínimos. A análise produz estimações dos parâmetros, isto é, $\hat{\beta}_s^e$ e também $\hat{\varepsilon}_s$. Além disso, estima um coeficiente de correlação múltipla $\mathbf{R_{es}^2}$.

Interpretação dos parâmetros

A partir dos parâmetros estimados $\hat{\beta}_s^e$, vamos estimar o ponto ideal, pois da equação (4.4.6) temos que:

$$\hat{\pi}_{sk} = -\frac{\hat{\beta}_{ks}}{2\hat{w}_s} \tag{4.4.9}$$

O coeficiente de correlação múltipla ($\mathbf{R_{es}^2}$) é usado como uma medida de ajuste do modelo. Também podemos testar a significância do coeficiente de correlação múltipla ($\mathbf{R_{es}^2}$) mediante as seguintes hipóteses:

$$H_0 : R_{es}^2 = 0$$
$$H_1 : R_{es}^2 > 0$$

A estatística F a ser utilizada é:

$$F = \frac{R_{es}^2}{\left(1 - R_{es}^2\right)} x \frac{(I - K - 2)}{(K+1)} \approx F_{[K+1],[I-K-2]}$$

Por outro lado, o modelo Ponto Ideal Simples ajustado aos dados é também usado para testar as seguintes hipóteses:

$$H_0 : R_{es}^2 = R_{vs}^2$$
$$H_1 : R_{es}^2 > R_{vs}^2$$

A estatística F a ser utilizada é:

$$F = \frac{\left(R_{es}^2 - R_{vs}^2\right)}{\left(1 - R_{es}^2\right)} x \frac{(I-K-2)}{1} \approx F_{[1],[I-K-2]}$$

A rejeição da hipótese H_0 implica que o modelo Ponto Ideal Simples avalia melhor a preferência do indivíduo s do que o modelo de Preferência Vetor. Além disso, também podemos dizer que a rejeição de H_0 indica que as preferências de cada indivíduo não são relacionadas linearmente para algumas dimensões (ou atributos).

- Estimação do Modelo Ponto Ideal Ponderado (P.I.P).

Já vimos que o modelo Ponto Ideal Ponderado é expresso da seguinte maneira (ver equação 4.3.4):

$$\delta_{is} = \sum_{k=1}^{K} w_{ks}\left(x_{ik} - \pi_{sk}\right)^2 + \varepsilon_s^* \qquad (4.4.10)$$

Desenvolvendo a equação (4.4.10) temos:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks} x_{ik}^2 + \sum_{k=1}^{K} \left(-2 w_{ks} x_{ik} \pi_{sk}\right) + \sum_{k=1}^{K} w_{ks} \pi_{sk}^2 + \varepsilon_s^* \qquad (4.4.11)$$

e, reparametrizando a equação (4.4.11) temos:

$$\varepsilon_s = \sum_{k=1}^{K} w_{ks} \pi_{sk} + \varepsilon_s^* \qquad (4.4.12)$$

$$\beta_{ks} = -2 w_{ks} \pi_{sk} \qquad (4.4.13)$$

Substituindo as equações (4.4.12) e (4.4.13) na equação (4.4.11) temos:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks} x_{ik}^2 + \sum_{k=1}^{K} \beta_{ks} x_{ik} + \varepsilon_s \qquad (4.4.14)$$

Na equação (4.4.14) se: $w_{ks} = w_s$, observamos que o modelo Ponto Ideal Ponderado reduz-se ao modelo Ponto ideal Simples. Analogamente, se $w_{ks} = 0$, então o modelo Ponto Ideal Ponderado reduz-se ao modelo de Preferência Vetor. Desta forma observamos que os modelos de preferência Vetor e Ponto Ideal Ponderado são casos particulares do modelo Ponto Ideal Ponderado.

Na forma matricial a equação (4.4.14) pode ser escrita da seguinte maneira:

$$\delta_S = X^W \beta_S^W + \varepsilon_S \tag{4.4.15}$$

onde:

$$\delta_S = \left(\delta_{1s} \quad \delta_{2s} \quad \cdots \quad \delta_{is} \quad \cdots \quad \delta_{Is}\right)^t_{Ix1}$$

$$X^W = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1K} & x_{11}^2 & x_{12}^2 & \cdots & x_{1K}^2 \\ x_{21} & x_{22} & \cdots & x_{2K} & x_{21}^2 & x_{22}^2 & \cdots & x_{2K}^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{I1} & x_{I2} & \cdots & x_{IK} & x_{I1}^2 & x_{I2}^2 & \cdots & x_{IK}^2 \end{pmatrix}_{Ix(K+K)}$$

$$\beta_S^W = \left(\beta_{1s} \quad \beta_{2s} \quad \cdots \quad \beta_{Ks} \quad w_{1s} \quad w_{2s} \quad \cdots \quad w_{Ks}\right)^t_{(K+K)x1}$$

A equação (4.4.15) pode ser interpretada como um modelo de Regressão Linear Múltipla, consequentemente, os parâmetros podem ser estimados via quadrados mínimos.

Interpretação dos parâmetros

A análise produz estimadores para os parâmetros, isto é, $\hat{\beta}_S^W$, e também $\hat{\varepsilon}_S$; além disso, estima um coeficiente de correlação múltipla R_{ws}^2. A partir dos parâmetros estimados $\hat{\beta}_S^W$, vamos estimar o ponto ideal, pois, da equação (4.4.13), temos que:

$$\hat{\pi}_{sk} = -\frac{\hat{\beta}_{ks}}{2\hat{w}_{ks}} \tag{4.4.16}$$

O coeficiente de correlação múltipla R_{ws}^2 é usado como uma medida de ajuste do modelo. Também podemos testar a significância do coeficiente de correlação múltipla R_{ws}^2 mediante as seguintes hipóteses:

$$H_0 : R_{ws}^2 = 0$$
$$H_1 : R_{ws}^2 > 0$$

A estatística F a ser utilizada é:

$$F = \frac{R_{ws}^2}{\left(1 - R_{ws}^2\right)} x \frac{(I - 2K - 1)}{(2K)} \approx F_{[2K],[I-2K-1]}$$

Por outro lado, o modelo Ponto Ideal Ponderado ajustado aos dados é também usado para testar as seguintes hipóteses:

$$H_0 : R_{ws}^2 = R_{es}^2$$
$$H_1 : R_{ws}^2 > R_{es}^2$$

A estatística F a ser utilizada é:

$$F = \frac{\left(R_{ws}^2 - R_{es}^2\right)}{\left(1 - R_{ws}^2\right)} x \frac{(I - 2K - 1)}{(K - 1)} \approx F_{[K-1],[I-2K-1]}$$

A rejeição de $\mathbf{H_0}$ indica que o modelo Ponto Ideal Ponderado é razoavelmente mais apropriado do que o modelo Ponto Ideal Simples. Isto, em termos práticos, quer dizer que os pesos para o indivíduo s não são iguais para todas as k dimensões (ou atributos).

Também pode ser feito o seguinte teste:

$$H_0 : R_{ws}^2 = R_{vs}^2$$
$$H_1 : R_{ws}^2 > R_{vs}^2$$

A estatística F a ser utilizada é:

$$F = \frac{\left(R_w^2 - R_v^2\right)}{\left(1 - R_w^2\right)} x \frac{(I - 2K - 1)}{K} \approx F_{[K],[I-2K-1]}$$

A rejeição da hipótese nula indica que o modelo Ponto Ideal Ponderado ajusta melhor os julgamentos do indivíduo s do que o modelo de Preferência Vetor . Além disso, pode-se afirmar que as preferências do indivíduo não são linearmente relacionadas para algumas dimensões (ou atributos).

- Estimação do Modelo Ponto Ideal Geral (P.I.G).

Demonstrou-se, na seção 4.3.2.3, que o modelo Ponto Ideal Geral pode ser escrito da seguinte maneira (ver equação 4.3.21):

$$\delta_{is} = \sum_{k=1}^{K} w_{ks} r_{kks} \left(x_{ik} - \pi_{sk}\right)^2 + \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \left(x_{ik} - \pi_{sk}\right)\left(x_{ik'} - \pi_{sk'}\right) + e_s^* \qquad (4.4.17)$$

Desenvolvendo o quadrado e multiplicando na equação (4.4.17) temos:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks} r_{kks} \left(x_{ik}^2 - 2x_{ik}\pi_{sk} + \pi_{sk}^2\right) + \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \left(x_{ik}x_{ik'} - 2x_{ik}\pi_{sk'} + \pi_{sk}\pi_{sk'}\right) + e_s^* \qquad (4.4.18)$$

e, escrevendo convenientemente a equação (4.4.18), temos:

$$\delta_{is} = \sum_{k=1}^{K} w_{ks} x_{ik}^2 + \sum_{k=1}^{K} \left(-2 w_{ks} x_{ik} \pi_{sk}\right) + \sum_{k=1}^{K} w_{ks} \pi_{sk}^2 +$$
$$+ \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} x_{ik} x_{ik'} +$$
$$+ \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \left(-2 x_{ik} \pi_{sk'}\right) + \tag{4.4.19}$$
$$+ \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \pi_{sk} \pi_{sk'} + \varepsilon_s^*$$

Também pode-se mostrar que:

$$\sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} x_{ik} x_{ik'} = \sum_{k=1}^{K} \sum_{(k' < k)}^{K} 2 w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} x_{ik} x_{ik'} \tag{4.4.20}$$

Seja a seguinte reparametrização:

$$\varepsilon_s = \sum_{k=1}^{K} w_{ks} \pi_{sk}^2 +$$
$$+ \sum_{k=1}^{K} \sum_{k'=1(k' \neq k)}^{K} w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \pi_{sk} \pi_{sk'} + e_s^* \tag{4.4.21}$$

$$b_{ks} = -2 w_{ks} \pi_{sk} + \sum_{k'=1(k' \neq k)}^{K} -2 w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \pi_{sk'} \tag{4.4.22}$$

$$b_{kk's} = 2 w_{ks}^{\frac{1}{2}} w_{k's}^{\frac{1}{2}} r_{kk's} \tag{4.4.23}$$

Para facilitar os cálculos, a equação (4.4.22) pode ser escrita na forma matricial:

$$b_S = -2M_S\pi_S \tag{4.4.24}$$

onde:

$$b_S = \begin{pmatrix} b_{1s} \\ b_{2s} \\ \vdots \\ \vdots \\ b_{Ks} \end{pmatrix}_{Kx1} \qquad \pi_S = \begin{pmatrix} \pi_{s1} \\ \pi_{s2} \\ \vdots \\ \vdots \\ \pi_{sK} \end{pmatrix}_{Kx1}$$

A matriz M_s é definida da seguinte maneira:

$$M_S = \begin{pmatrix} m_{11s} & m_{12s} & \cdots & m_{1Ks} \\ m_{21s} & m_{22s} & \cdots & m_{2Ks} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1s} & m_{K2s} & \cdots & m_{KKs} \end{pmatrix}_{KxK}$$

onde cada elemento da matriz M_s é escrito como:

$$m_{kk's} = \begin{cases} w_{ks}^{\frac{1}{2}} \; w_{k's}^{\frac{1}{2}} \; r_{kk's} \rightarrow \text{se: } k \neq k' \\ w_{ks} \rightarrow \text{se: } k = k' \end{cases} \qquad \forall\ k,k'=1,2,....,K$$

Substituindo as equações (4.4.20), (4.4.21), (4.4.22), (4.4.23) na equação (4.4.19) obtemos a seguinte equação:

$$\delta_{is} = \sum_{k=1}^{K} b_{ks}x_{ik} + \sum_{k=1}^{K} w_{ks}x_{ik}^2 + \sum_{k=1}^{K}\sum_{(k'<k)}^{K} b_{kk's}x_{ik}x_{ik'} + \varepsilon_s \tag{4.4.25}$$

que pode ser escrita na forma matricial como:

$$\delta_S = X^g\beta_S^g + \varepsilon_S \tag{4.4.26}$$

onde:

$$\delta_s = \left(\delta_{1s} \quad \delta_{2s} \quad \cdots \quad \delta_{is} \quad \cdots \quad \delta_{Is}\right)^t_{Ix1}$$

$$X^g = \begin{pmatrix} x_{11} & \cdots & x_{1K} & x_{11}^2 & \cdots & x_{1K}^2 & x_{11}x_{12} & \cdots & x_{11}x_{1K} \\ x_{21} & \cdots & x_{2K} & x_{21}^2 & \cdots & x_{2K}^2 & x_{21}x_{22} & \cdots & x_{21}x_{2K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{I1} & \cdots & x_{IK} & x_{I1}^2 & \cdots & x_{IK}^2 & x_{I1}x_{I2} & \cdots & x_{I1}x_{IK} \end{pmatrix}_{Ix\left(2K+\frac{1}{2}K(K-1)\right)}$$

$$\beta_s^g = \left(b_{1s} \quad \cdots \quad b_{Ks} \quad w_{1s} \quad \cdots \quad w_{Ks} \quad b_{21s} \quad \cdots \quad b_{K(K-1)s}\right)^t_{\left(2K+\frac{1}{2}K(K-1)\right)}$$

A equação (4.4.26) pode ser interpretada como um modelo de Regressão Linear Múltipla, implicando que os parâmetros podem ser estimados via quadrados mínimos.

Interpretação dos parâmetros

A análise produz estimativas dos parâmetros, as quais são denotadas por, $\hat{\beta}_s^g$, e também $\hat{\varepsilon}_s$; além disso, estima um coeficiente de correlação múltipla $\mathbf{R}_{gs}^2$.

A partir dos parâmetros estimados $\hat{\beta}_s^g$, podemos estimar os elementos da matriz $\mathbf{M}_s$, isto é, da equação (4.4.23). Primeiramente, estimamos $\hat{r}_{kk's}$ ou seja:

$$\hat{r}_{kk's} = \frac{\hat{b}_{kk's}}{2\hat{w}_{ks}^{\frac{1}{2}}\,\hat{w}_{k's}^{\frac{1}{2}}} \qquad (4.4.27)$$

Depois, usando a equação (4.4.27), podemos estimar facilmente a matriz $\hat{M}_S$, para então estimarmos o ponto ideal para cada indivíduo através da equação (4.4.24), ou seja:

$$\hat{\pi}_S = -\frac{\hat{b}_S}{2\hat{M}_S} = -\frac{1}{2}\hat{b}_S\hat{M}_S^{-1} \qquad (4.4.28)$$

O coeficiente de correlação múltipla $\mathbf{R_{gs}}^2$ é usado como uma medida de ajuste do modelo. Também podemos testar a significância do coeficiente de correlação múltipla $\mathbf{R_{gs}}^2$ mediante as seguintes hipóteses:

$$H_0 : R_{gs}^2 = 0$$
$$H_1 : R_{gs}^2 > 0$$

A estatística F a ser utilizada é:

$$F = \frac{R_{gs}^2}{\left(1-R_{gs}^2\right)} x \frac{\left[I-\frac{1}{2}K(K+3)-1\right]}{\left[\frac{1}{2}K(K+3)\right]} \approx F_{\left[\frac{1}{2}K(K+3)\right],\left[I-\frac{1}{2}K(K+3)-1\right]}$$

Por outro lado, o modelo Ponto Ideal Ponderado ajustado aos dados é também usado para testar as seguintes hipóteses:

$$H_0 : R_{gs}^2 = R_{ws}^2$$
$$H_1 : R_{gs}^2 > R_{ws}^2$$

A estatística F a ser utilizada é:

$$F = \frac{\left(R_{gs}^2 - R_{ws}^2\right)}{\left(1-R_{gs}^2\right)} x \frac{\left[I-\frac{1}{2}K(K+3)-1\right]}{\left[\frac{1}{2}K(K-1)\right]} \approx F_{\left[\frac{1}{2}K(K-1)\right],\left[I-\frac{1}{2}K(K+3)-1\right]}$$

A rejeição de $\mathbf{H_0}$ sugere que o termo da interação $r_{kk's}$ não é igual a zero para um o mais pares de dimensões (ou atributos).

A segunda estatística F usada para testar as hipóteses:

$$H_0 : R_{gs}^2 = R_{es}^2$$
$$H_1 : R_{gs}^2 > R_{es}^2$$

é dada por:

$$F = \frac{\left(R_{gs}^2 - R_{es}^2\right)}{\left(1 - R_{gs}^2\right)} x \frac{\left[I - \frac{1}{2}K(K+3) - 1\right]}{\left[\frac{1}{2}K(K+1) - 1\right]} \approx F_{\left[\frac{1}{2}K(K+1)-1\right],\left[I-\frac{1}{2}K(K+3)-1\right]}$$

A rejeição da hipótese nula indica que o modelo Ponto Ideal Geral é mais apropriada para avaliar os julgamentos do indivíduo s do que o modelo Ponto Ideal Simples. Isto estaria indicando que não se presume igual ponderação para todas as dimensões (ou atributos). Outro motivo é a presença de interação entre algumas dimensões (ou atributos).

Também pode-se testar as seguintes hipóteses:

$$H_0 : R_{gs}^2 = R_{vs}^2$$
$$H_1 : R_{gs}^2 > R_{vs}^2$$

A estatística F usada é :

$$F = \frac{\left(R_{gs}^2 - R_{vs}^2\right)}{\left(1 - R_{gs}^2\right)} x \frac{\left[I - \frac{1}{2}K(K+3) - 1\right]}{\left[\frac{1}{2}K(K+1)\right]} \approx F_{\left[\frac{1}{2}K(K+1)\right],\left[I-\frac{1}{2}K(K+3)-1\right]}$$

A rejeição da hipótese nula indica que o modelo Ponto Ideal Geral é mais apropriada para avaliar os julgamentos do indivíduo s do que o modelo de Preferência Vetor. Outro motivo para a rejeição é a possivel existência de interação entre algumas par de dimensões (ou atributos). Também a rejeição da hipótese nula estaria indicando a não linearidade entre as preferências do indivíduo e as dimensões (ou atributos).

4.4.2. Estimação Não Métrica

A estimação não métrica, será usada quando as avaliações dos indivíduos (δ_{is}) forem medidas em escala ordinal. A idéia básica é encontrar para cada indivíduo uma transformação monótona, que aplicada a δ_{is} faça com que os modelos já mencionados possam ser ajustados via quadrados mínimos.

A seguir mencionamos o algoritmo descrito por Artes (1991) para estimar os modelos nesta parte não métrica.

4.4.2.1.- Algoritmo da estimação não métrica.

- **Passo 1**

Ordenar as avaliações dos indivíduos δ_{is} de forma ascendente.

- **Passo2**

Estimar o modelo utilizando δ_{is} como variável dependente.

- **Passo 3**

Com o modelo estimado calcular os valores ajustados para cada estímulo.

- **Passo 4**

Aplicar o algoritmo de Regressão Isotônica (ver apêndice) aos valores ajustados pelo modelo, o qual vamos denotar por $y^{(1)}$.

- **Passo 5**

Estimar novamente o modelo utilizando $y^{(1)}$ como variável dependente.

- **Passo 6**

Repetir os passos 2, 3 e 4 até que algum critério de convergência seja satisfeito.

Ao final do processo serão obtidas as estimativas apropriadas e os coeficientes de regressão para serem utilizados na estimação dos pontos ideais, matrizes de rotação e pesos.

Os testes F mencionados na parte métrica perdem sua validade, mas podem ser utilizados como indicadores da melhoria do ajuste ao se mudar de modelo. Por outro lado, sugere fazer testes não paramétricos para medir a qualidade do ajuste do modelo.

4.4.2.1.- Critério de Convergência

Como critério de parada do processo iterativo, Artes (1991) sugere, calcular a diferença relativa máxima entre os parâmetros estimados em iterações sucessivas:

$$DIFMAX = \max\left\{\frac{\beta_k^{(j)} - \beta_k^{(j-1)}}{\beta_k^{(j-1)}}; \forall k = 1,2,.....\right\}$$

Quando por duas iterações sucessivas, essa diferença foi inferior a 0.01, interrompe-se o processo de estimação.

4.5.- Observações a respeito dos Modelos de Preferência

Como em todo modelo estatístico, existe sempre críticas construtivas sobre o modelo. Isto é fácil de notar quando trabalhamos, sobre tudo, na prática, pois, às vezes, os dados a serem analisados não satisfazem as suposições teóricas.

Por exemplo, nem sempre a estimativa do suposto Ponto Ideal pode ser interpretada como a localização do estímulo de maior preferência. Isto só acontece se duas condições forem satisfeitas:

a.- A estimativa $\hat{w}_{kS}$ do peso correspondente ao atributo k do indivíduo S é positiva.
b.- Que por ocasião do planejamento se estabeleça que os estímulos sejam avaliados de modo que δ_{is} cresça conforme diminua sua preferência.

Davison (1983) assinala que Caroll (1972) discorre sobre a violação da condição a. O autor sugere que quando esta for violada, $\hat{\pi}_{kS}$ seja interpretado como a localização do estímulo de menor preferência, denominando este ponto como Ponto Anti-ideal. Deste modo, quanto mais próximo se estiver deste ponto, menor será a preferência. Na prática também costuma-se mudar de sinal para os pontos ideais quando estes tiverem pesos negativos.

Capítulo V

APLICAÇÃO EM MERCADOLOGIA

5.1. Introdução ao problema

No cenário financeiro atual estão surgindo bancos com novos conceitos. É cada vez mais freqüente que um banco assuma operações de outros, e é crescente o clima de competitividade na busca de novos clientes e manutenção dos atuais. Neste contexto, a pesquisa de mercado tem papel cada vez mais relevante, uma vez que a rentabilidade repousa em sua carteira de clientes e não mais nas aplicações financeiras, como ocorria nos tempos de inflação alta. Recentemente, uma empresa de pesquisa de mercado realizou uma pesquisa quantitativa, visando o monitoramento e avaliação do atendimento prestado ao novo segmento de clientes. Por razões de ética, não serão revelados os nomes das empresas envolvidas. Ressaltamos, no entanto, que se trata de um caso real que tem o propósito de ilustrar o potencial e as limitações das técnicas de Escalonamento Multidimensional e Modelos de Preferência, desenvolvidos anteriormente.

Através do desenvolvimento desta aplicação, além da ilustração, fica delineada uma estratégia de análise e sua fácil operacionalização, de forma que a mesma possa ser utilizada também em outras áreas, sem maiores dificuldades. Diante do exposto, optamos por documentar, neste trabalho, a análise de apenas um dos aspectos pesquisados: o aspecto relativo à avaliação de imagem institucional, tal como percebida através de um item do instrumento (questionário) utilizado na coleta dos dados.

Este capítulo é composto de 9 seções. Na seção 5.2, é descrito o processo de coleta de dados, bem como a escolha dos estímulos estudados. Na seção 5.3 é apresentada a metodologia estatística adotada. Na seção 5.4 são apresentados os resultados para cada diretoria. Na seção 5.5 são apresentados os resultados para o total, e na seção 5.6 os resultados para estrutura de grupos. Na seção 5.7 é feita a análise comparativa dos resultados via MDS. Nas seções 5.8 e 5.9 focalizamos o desenvolvimento metodológico dos Modelos de Preferência, bem como a

discussão final dos resultados. Na última seção, 5.10, apresentamos as conclusões gerais da aplicação.

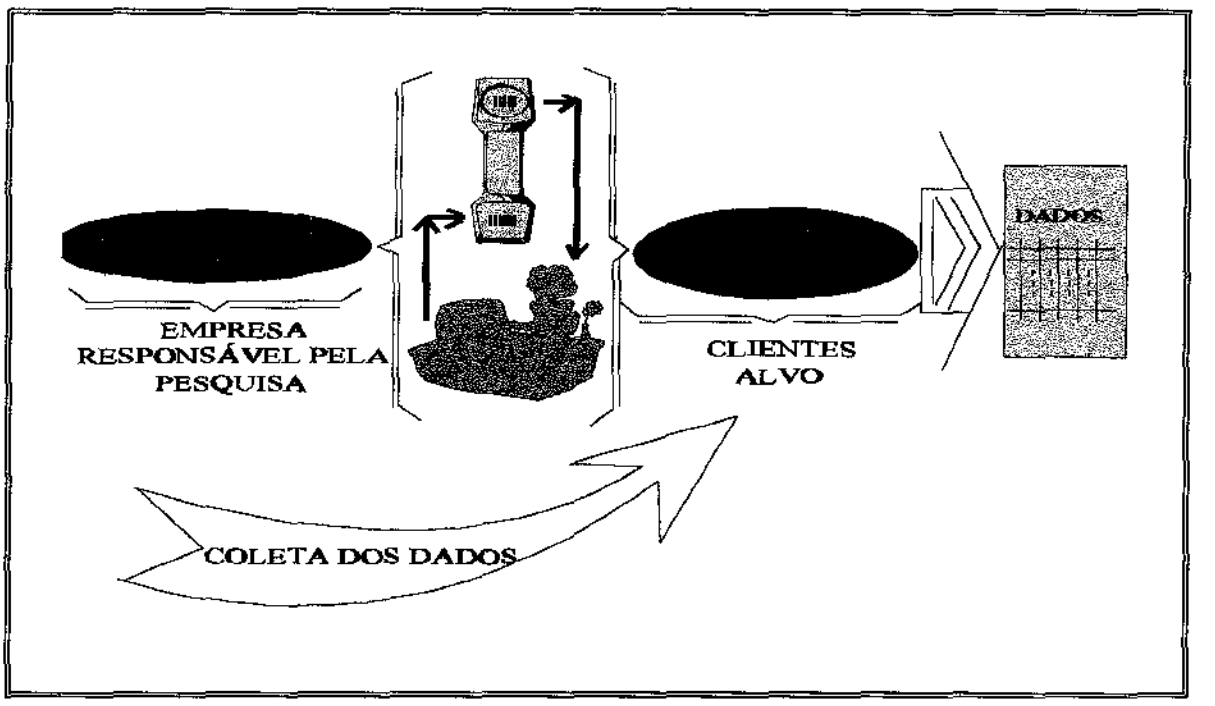
5.2. Coleta de Dados.

5.2.1. Metodologia

A população alvo, neste caso, foi definida como clientes PJ (pessoas jurídicas), que compõem o segmento de interesse da instituição financeira, recentemente incorporado.

A metodologia de coleta adotada foi a de entrevistas telefônicas, sendo o instrumento um questionário estruturado. Para ilustrarmos melhor como ocorreu a coleta de dados, mostramos, a seguir, o seguinte esquema:

Gráfico 5.1. Esquema da Coleta de dados.



O interesse da avaliação institucional foi contemplado em um item do questionário, no qual as variáveis não foram explicitamente mensuradas, mas coletadas indiretamente nas comparações intercaracterísticas. Assim, aos respondentes foi solicitado que ordenassem, de

forma crescente em importância, características institucionais de bancos entre sete características selecionadas. Na linguagem mercadológica, estas são chamadas de estímulos.

5.2.2. Amostra

O universo abordado é o de empresas que compõem o "Segmento Corporate" (clientes de maior importância). Deverão responder por estas empresas, o seu Diretor Financeiro ou Responsável por contas a pagar nas praças de Rio Grande do Sul, São Paulo, Rio de Janeiro, Paraná, Santa Catarina, Ribeirão Preto, Campinas, Nordeste, Minas Gerais, Brasília, Goiás, Espírito Santo, German Desk e Japan Desk.

O tamanho da amostra foi de 513 indivíduos, divididos em 12 regiões, os quais foram denominados de diretorias. A seguir mostramos a tabela 5.1 descrevendo as diretorias, regiões e tamanho de cada diretoria.

Tabela 5.1 : Estrutura dos dados por diretorias.

DIRETORIAS	INDIVÍDUOS	REGIÕES
diretoria 1	$n_1 = 49$	Rio Grande do Sul
diretoria 2	$n_2 = 49$	São Paulo
diretoria 3	$n_3 = 47$	São Paulo
diretoria 4	$n_4 = 43$	São Paulo
diretoria 5	$n_5 = 47$	Rio de Janeiro
diretoria 6	$n_6 = 49$	Paraná / Sta. Catarina
diretoria 7	$n_7 = 51$	Riberão Preto
diretoria 8	$n_8 = 52$	Campinas
diretoria 9	$n_9 = 50$	Norte / Nordeste
diretoria 10	$n_{10} = 51$	Minas Gerais / Brasília / Goias/ Espírito Santo
diretoria 11	$n_{11} = 12$	German Desk
diretoria 12	$n_{12} = 13$	Japan Desk

5.2.3. Descrição dos Estímulos

Os estímulos que mencionaremos a seguir são referentes às características dos bancos, ou seja, aquelas que indicam aspectos associados à instituição como um todo.

- Estímulo 1 : Localização geográfica das agências.
- Estímulo 2 : Especialização em grandes clientes.

- Estímulo 3 : Solidez da instituição.
- Estímulo 4 : Liderança de mercado.
- Estímulo 5 : Agilidade operacional da estrutura administrativa.
- Estímulo 6 : Autonomia de profissionais.
- Estímulo 7 : Parceria no relacionamento com empresas.

A pergunta formulada para o indivíduo foi: Qual destes estímulos é, na sua opinião, o mais importante em $1^{\underline{o}}$ lugar? em $2^{\underline{o}}$ lugar?,..., em $7^{\underline{o}}$ lugar?. Pode-se notar que cada indivíduo estava sujeito a dar sua preferência, de mais importante a menos importante, para os sete estímulos.

5.2.4. Criação do arquivo

Os dados foram digitados em arquivos ASCII e posteriormente lidos pelo softaware estatística SAS, através do qual foi gerado um único arquivo de dados endereçado pela variável classificadora diretoria, e 7 estímulos (EST1-EST7). O arquivo em questão continha 513 indivíduos, 7 estímulos e uma variável classificadora. Este arquivo foi básico para o processamento dos dados até a identificação das novas dimensões (ou atributos), bem como para os Pontos Ideais.

A estrutura deste arquivo pode ser vista na tabela 5.2, onde cada linha corresponde à avaliação de um indivíduo e suas respectivas preferências, e as colunas representam os 7 estímulos, bem como a variável classificadora. Percebe-se que o indivíduo 1 pertence à diretoria 2 (grupo 2). Por exemplo, para este indivíduo, o estímulo 5 era o mais importante, seguido pelos estímulos 6, 1, e assim sucessivamente o estímulo 4 foi o menos importante.

Tabela 5.2 . Estrutura do arquivo de dados em forma inicial.

Indivíduos	Diretoria	EST1	EST2	EST3	EST4	EST5	EST6	EST7
ind.1	dir2	3	4	5	7	1	2	6
ind.2	dir4	7	6	1	5	3	2	4
ind.3	dir1	7	1	2	5	6	4	3
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
ind.513	dir3	7	5	4	6	2	1	3

5.3. Metodologia Estatística

O processo analisa a percepção dos indivíduos, na moldura amostral já descrita, quanto às preferências dos características institucionais do banco.

A metodologia estatística adequada para este tipo de pergunta pode ser estruturada em duas partes:

i.- A primeira parte consiste da aplicação do Escalonamento Multidimensional (MDS) para obter uma representação gráfica dos estímulos de forma a visualizar suas posições relativas, na menor dimensão possível.

ii.- A segunda parte consiste na estimação dos Pontos Ideais individuais através dos Modelos de Preferência, de forma a obter o comportamento dos indivíduos frente aos estímulos.

Em ambas as partes, como se poderá observar, a metodologia usada foi o MDS e os Modelos de Preferência. O esquema apresentado na figura 5.2 documenta as diversas etapas usadas no procedimento de Modelos de Preferência via MDS. Este esquema de metadologia estatística dos dados, apresentado na figura 5.2, descreve como foram analisados os dados. Além disso, descrevemos em cada etapa, os procedimentos do S.A.S usados.

Os procedimentos do SAS usados neste trabalho foram os seguintes:

- Procedimento **FREQ**. ⇨ Procedimento usado para calcular tabelas de frequências.

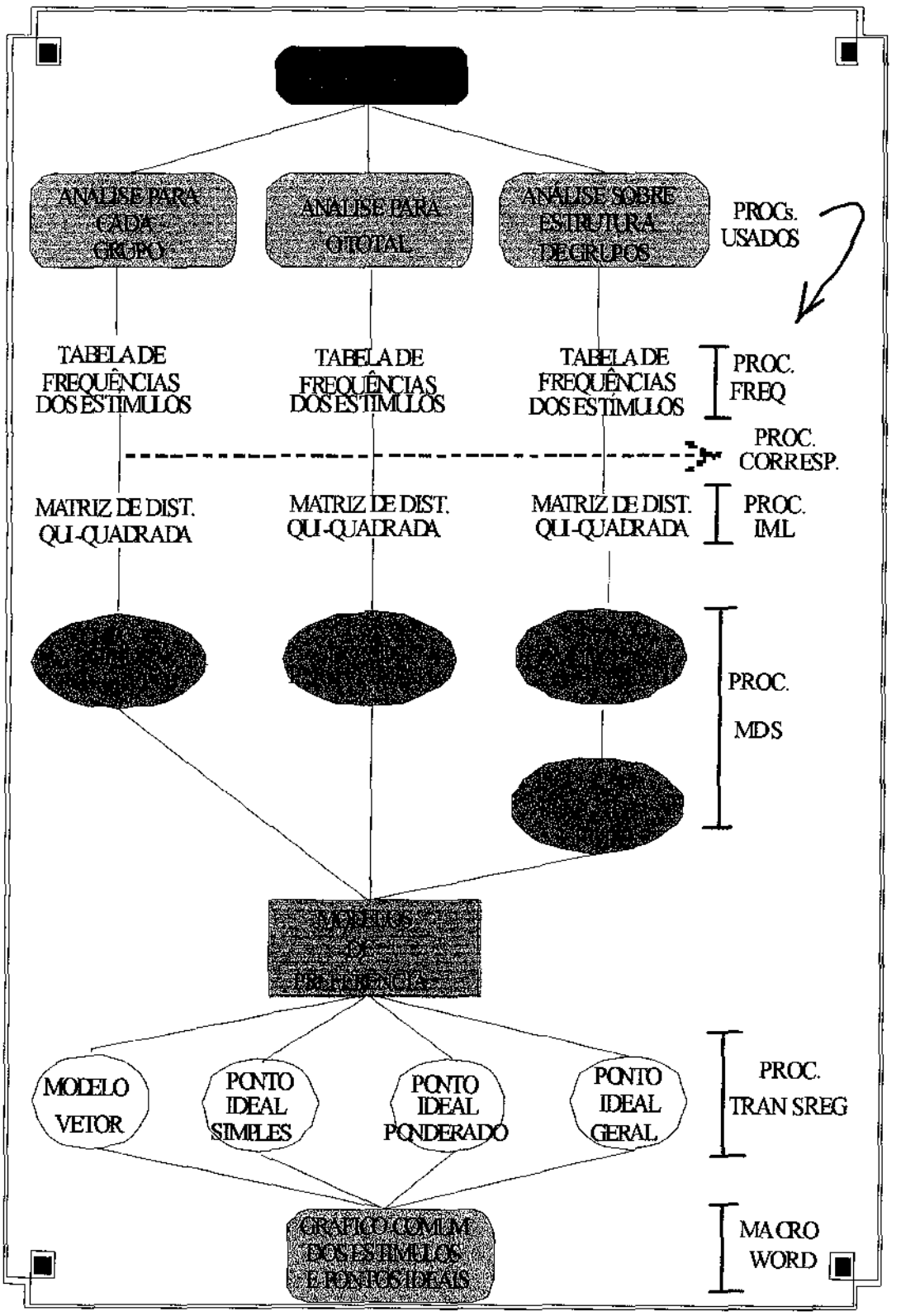
- Procedimento **CORRESP**. ⇨ Este procedimento permite realizar a técnica da Análise de Correspondência.

- Procedimento **IML**. ⇨ Procedimento usado para calcular distâncias.

- Procedimento **MDS**. ⇨ Este procedimento permite realizar o Escalonamento Multidimensional.

- Procedimento **TRANSREG**. ⇨ Este procedimento permite realizar os Modelos de Preferência.

Figura 5.2. Esquema da Análise Estatística.



5.4. Resultados via MDS para Cada Grupo (Diretoria)

A solução foi obtida utilizando-se a técnica de Escalonamento Multidimensional (MDS) não métrico e MDS sobre estrutura de grupos. Para este propósito, o primeiro passo foi encontrar uma matriz de distâncias que caracterizasse o conjunto de dados, tal como foi visto nos capítulos II e III do presente trabalho. Neste caso, como estamos diante de dados de nível ordinal, a primeira tarefa foi reagrupar os dados em freqüências segundo as respostas dos indivíduos, para podermos, depois, achar a distância Qui-quadrado (ver apêndice A) entre os estímulos. Optamos por esta medida de distância porque a distância euclidiana não é aconselhável para dados ordinais.

Neste seção, tratamos cada diretoria separadamente, isto é, não estamos considerando estrutura de grupos. Para o desenvolvimento desta seção, foram usados os passos que detalhamos a seguir para a diretoria 1. As demais diretorias foram tratadas analogamente.

Diretoria 1

- Passo 1.

O primeiro passo foi agrupar os indivíduos em sua respectiva diretoria. Para isto, usamos a informação da tabela 5.2, utilizando-se o programa 1 (ver apêndice G). Assim, a tabela 5.3, mostra as avaliações dos indivíduos da diretoria 1.

Tabela 5.3. Avaliações dos indivíduos da Diretoria 1.

Indivíduos	Diretoria	EST1	EST2	EST3	EST4	EST5	EST6	EST7
ind. 1	dir 1	7	1	2	5	6	4	3
ind. 2	dir 1	7	5	4	6	1	3	2
ind. 3	dir 1	4	6	1	7	3	5	2
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
ind. 49	dir 1	7	4	6	5	2	3	1

- Passo 2.

Os dados da tabela 5.3, pertencentes à diretoria 1, foram arrumados em uma nova tabela. Esta nova tabela é uma tabela de contingência, onde as colunas pertencem aos estímulos

e as linhas pertencem às freqüências das preferências. Por exemplo, o elemento da tabela 5.4, correspondente a coluna EST1 e a linha R1 é o número 3. Este número está indicando que, da Diretroria 1, que contém 49 indivíduos, 3 deles avaliaram o Estímulo 1 como o mais importante. Para este passo foi usado o programa 2 (ver apêndice G), no qual basicamente foi usado o Procedimento FREQ. do SAS.

Tabela 5.4. Tabela de Contingência para Diretoria 1.

	EST1	EST2	EST3	EST4	EST5	EST6	EST7	$n_{i.}$	$n_{i.}/n$
R1	3	4	20	0	2	6	14	49	0.1429
R2	3	2	11	1	7	10	15	49	0.1429
R3	3	7	5	6	13	9	6	49	0.1429
R4	2	8	8	3	11	8	9	49	0.1429
R5	7	7	2	13	8	10	2	49	0.1429
R6	8	9	3	16	8	3	2	49	0.1429
R7	23	12	0	10	0	3	1	49	0.1429
$n_{.j}$	49	49	49	49	49	49	49	343	0.1429
$n_{.j}/n$	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429		

Nota-se que as marginais das linhas e colunas (parte sombreada), são as mesmas, devido à forma como se obtiveram os dados. Neste caso, o método adotado foi o ordenamento através dos postos (ver capítulo 1).

- <u>Passo 3</u>.

Agora, trataremos de encontrar uma medida de dissimilaridade entre os estímulos. A princípio, pensou-se na distância Euclidiana, mas a análise com a distância Euclidiana não apresentou bons resultados. Então optou-se por procurar outra medida de dissimilaridade. Além disso, os dados encontravam-se na escala ordinal. Diante deste problema, optamos por procurar alguns coeficientes de similaridade, mencionados no apêndice A, os quais também não proporcionaram resultados favoráveis.

Depois de pesquisar, optamos então por trabalhar com a distância Qui-quadarada, pois era a medida mais recomendável para este tipo de dados. A fórmula usada é a mencionada no apêndice A, equação (a.30). Assim a distância Qui-quadrada entre o estímulo 1 e o estímulo 2 é: 0.5130. Assim sucessivamente calcularam-se as 21 distâncias, as quais foram arrumadas na

matriz 5.1, denominada de matriz de distâncias. Para isto, foi usado o programa 3 (ver apêndice G), empregando-se basicamente o Procedimento IML do SAS.

Matriz 5.1. Matriz de distância Chi-quadrada entre estímulos para Diretoria 1.

$$D_{QUI-Q} = \begin{array}{c} \\ EST1 \\ EST2 \\ EST3 \\ EST4 \\ EST5 \\ EST6 \\ EST7 \end{array} \begin{pmatrix} 0 & & & & & & \\ 0.5130 & 0 & & & & & \\ 2.8330 & 1.5914 & 0 & & & & \\ 0.8511 & 0.3847 & 2.6698 & 0 & & & \\ 2.1218 & 0.6412 & 1.3815 & 0.9968 & 0 & & \\ 1.6438 & 1.5771 & 0.8336 & 1.1017 & 0.2565 & 0 & \\ 2.5299 & 1.3582 & 0.1632 & 2.4074 & 0.9735 & 0.4896 & 0 \end{pmatrix}_{7x7}$$

- **Passo 4.**

Dado que já encontramos a medida de similaridade entre os estímulos, agora é preciso encontrar uma configuração de pontos (menor dimensão possível) que possa representar o comportamento dos estímulos. Para isto foi usada a técnica do MDS não métrico. Paralelamente, foi usado o programa 4 (ver apêndice G), em que trabalhamos com o Procedimento MDS do SAS.

Tabela 5.5. Solução do MDS Não Métrico para diretoria 1

ESTÍMULOS	dimensão 1	dimensão 2
EST1	1.73514	0.92212
EST2	0.68235	0.23358
EST3	-1.79318	0.46527
EST4	1.52092	-0.60518
EST5	-0.25305	-0.98792
EST6	-0.43274	-0.28762
EST7	-1.45943	0.25976

A medida do bondade de ajuste chamada de STRESS foi : 0.00390, a qual está indicando que o ajuste da configuração dos estímulos foi muito bom. Por outro lado, a

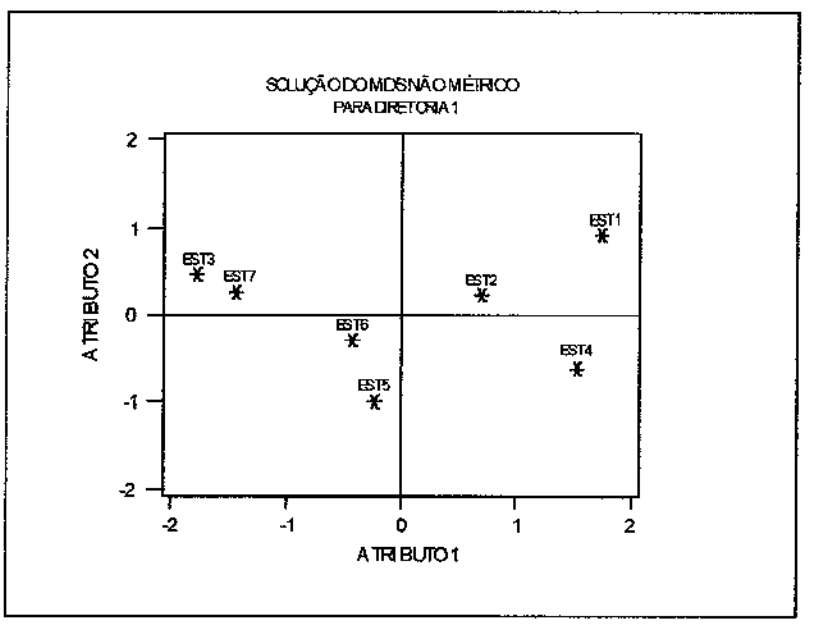
representação obtida dos sete estímulos no plano preserva ao máximo, a ordem das distâncias da matriz 5.1.

Nota-se que o ajuste do MDS foi feito em duas dimensões. Este ajuste também foi feito para três dimensões, onde a medida de ajuste, o STRESS, foi 0.00125. Percebe-se que a medida não melhorou, então decidimos trabalhar em duas dimensões. Também é de se esperar que estas novas dimensões gerem os novos atributos, os quais servirão para ajustar os modelos de preferência.

- Passo 5.

Neste passo, representamos graficamente a solução do MDS não Métrico, dada pelo passo 4. Deste modo, poderemos observar o comportamento dos estímulos em um plano bi-dimensional, e visualizar graficamente os estímulos. O gráfico foi confeccionado no Macro Word.

Gráfico 5.3. Comportamento dos estímulos da diretoria 1



STRESS = 0.00125

Como se observa no gráfico 5.3, os estímulos 3 e 7 estão em contra-posição aos estímulos 1 e 4. Além disso, os estímulos 3 e 7 estão bastante próximos, indicando que esses estímulos têm algo em comum. O mesmo acontece com os estímulos 6 e 5.

A seguir no passo 6 será utilizado a técnica de **Análise de Correspondência** e neste sentido faremos uma breve menção à esta técnica pois não é objetivo deste trabalho desenvolve-la.

A análise de Correspondência, é uma técnica multivariada descritiva para análise exploratória de dados categorizados. Ela converte uma matriz de dados categorizados (como exemplificado na tabela 5.4), em um particular tipo de gráfico que exibe as linhas e colunas da matriz como pontos de um espaço vetorial de dimensão menor que a original, de maneira que as relações entre as linhas, colunas e entre linhas e colunas conjuntamente possam ser interpretadas.

O desenvolvimento teórico da Análise de Correspondência, esta baseada sobre dois aspectos, geométrico e algébrico, a qual fazem com que ela pertença a uma família de técnicas de disposição gráfica que são baseadas em aproximação de matrizes, por outra de posto menor por meio da descomposição em valore singulares (ver apêndice C).

Assim o problema da Análise de Correspondência passa a ser o de estimar um subespaço que melhor ajuste a nuvem de pontos no espaço euclidiano. Este ajuste é feito pelo método de quadrado mínimo ponderado onde é utilizada a distância euclidiana ponderada (distância Qui-quadrada) e a massa de pontos como seu peso.

A abordagem da Análise de Correspondência enfatiza a representação geométrica, revelando a estrutura dos dados de forma ótima e sem necessidade de suposição de modelo de distribuição para os dados.

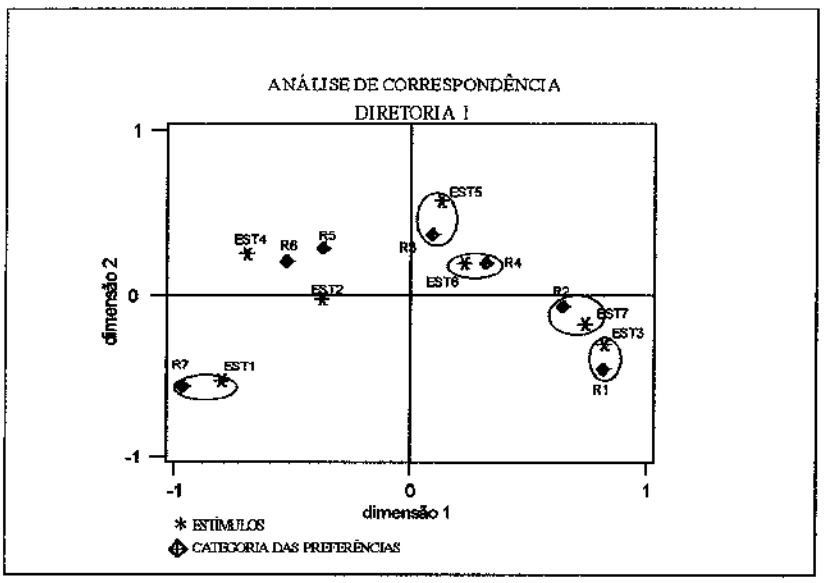
Podemos então dizer que um dos propósitos da Análise de Correspondência é o de reduzir dimensão do espaço vetorial, conservando da melhor forma possível a configuração inicial. Isto significa que podemos considerar a Análise de Correspondência como sendo um algoritmo de redução de dados que fornece imagens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

Neste contexto faremos a aplicação da Análise de Correspondência para nosso caso.

- **Passo 6.**

A seguir, utilizou-se a Análise de Correspondência para cada diretoria, com a finalidade de poder comparar os resultados com a solução do MDS. Além disso, é recomendável usar a Análise de Correspondência como uma técnica complementar quando os dados estiverem na escala ordinal e puderem ser rearranjados em tabelas de frequências. Este resultado foi obtido usando-se o programa 5 (ver apêndice G). Basicamente, o nosso interesse estava em estudar o perfil linha e coluna conjuntamente, isto é, saber o comportamento conjunto dos estímulos e as frequências das preferências. A seguir mostramos o gráfico da solução da Análise de Correspondência, feito no Macro Word.

Gráfico 5.4. Comportameno conjunto dos estímulos e as frequências das preferências Diretoria 1



% DE INÉRCIA = 92.75%

A inércia das duas dimensões foi de 92.75%, a qual estaria indicando um bom ajuste dado pelas duas dimensões. Como pode-se observar no gráfico 5.4, o estímulo 3 está junto a r1, o qual está indicando que o estímulo 3 foi o mais preferido pelos indivíduos. Verifica-se também que o menos preferido foi o estímulo 1.

Comparando o gráfico dado pelo MDS e o gráfico dado pela Análise de Correspondência, percebe-se que um é reflexo do outro, isto é, os estímulos têm quase o mesmo comportamento. Também percebe-se que os estímulos 7 e 3 estão em contra-posição

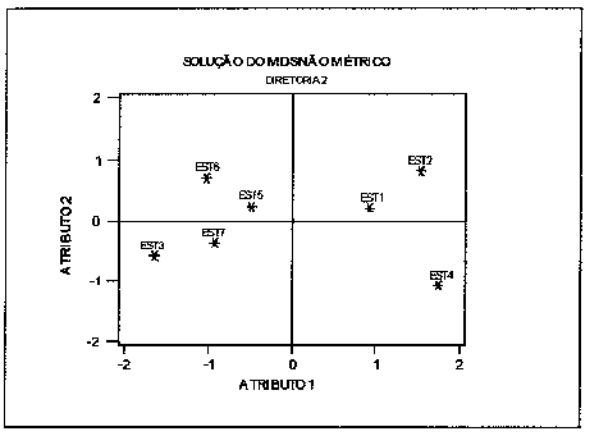
aos estímulos 4 e 1, e também os estímulos 3 e 7 novamente se encontram muito próximos. O mesmo acontece com os estímulos 5 e 6.

Procedendo de modo análogo para as 12 diretorias, mostramos a seguir, de forma resumida, os gráficos construídos a partir da solução do MDS para cada diretoria. Também apresentamos a solução dada pela Análise de Correspondência, assim como a medida de ajuste para o MDS e o total de inércia explicada pela solução da Análise de Correspondência.

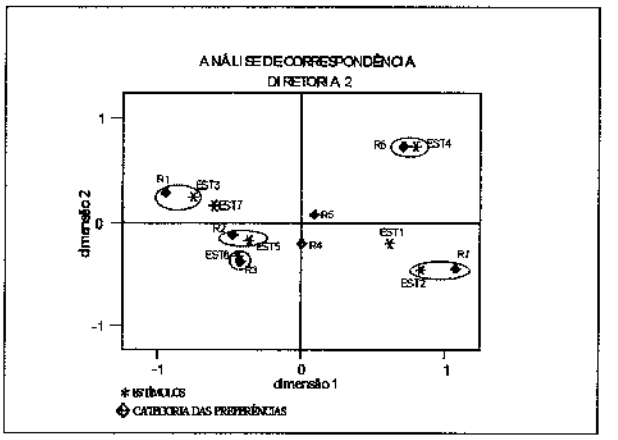
Gráfico 5.5. Síntese dos resultados obtidos para cada diretoria.

MDS NÃO MÉTRICO | **ANÁLISE DE CORRESPONDÊNCIA**

Diretoria 2

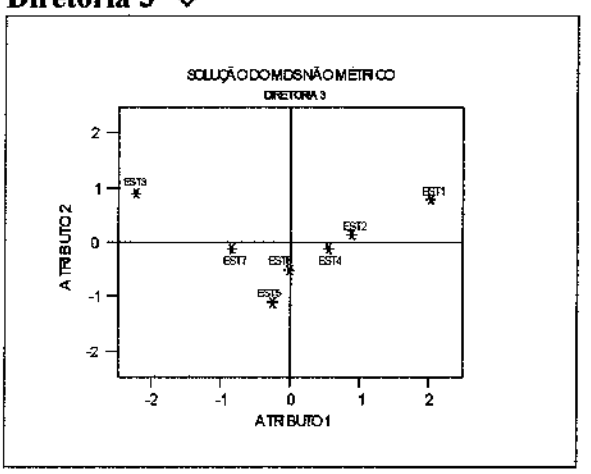


STRESS = 0.01968

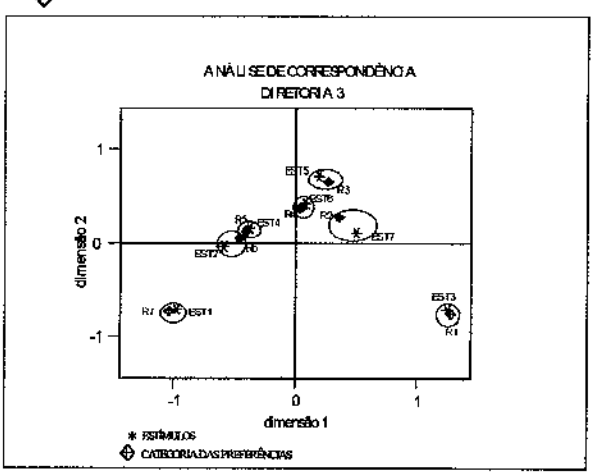


% INÉRCIA = 78.62 %

Diretoria 3

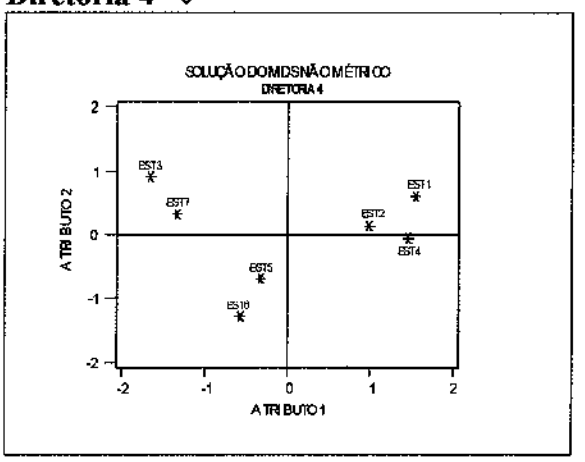


STRESS = 0.01968

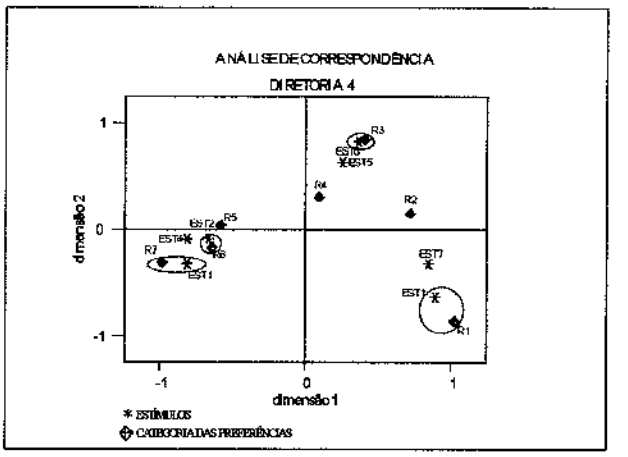


% INÉRCIA = 78.62 %

Diretoria 4

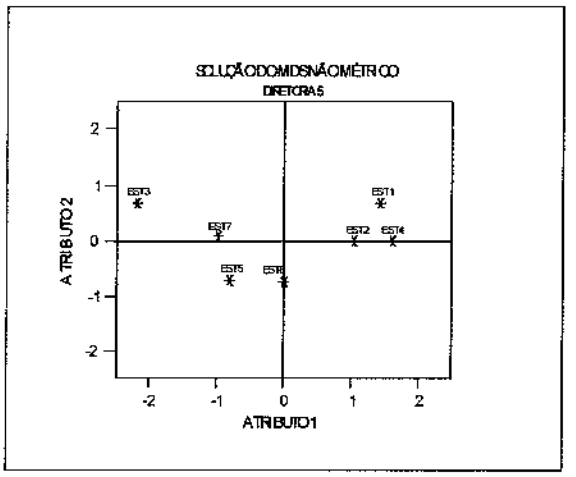


STRESS = 0.01009

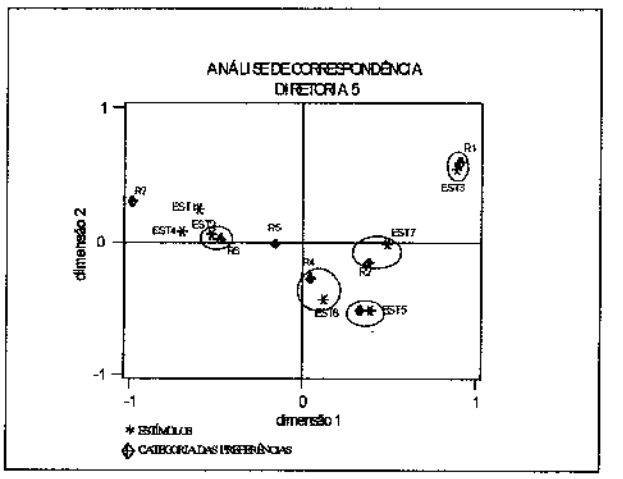


% INÉRCIA = 91.80 %

Diretoria 5

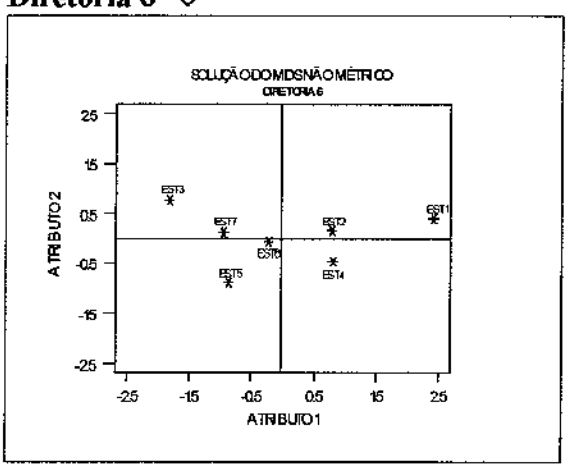


STRESS = 0.01424

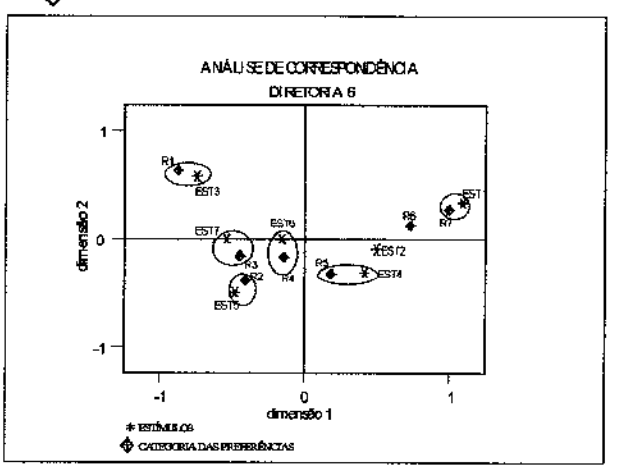


% INÉRCIA = 85.36 %

Diretoria 6

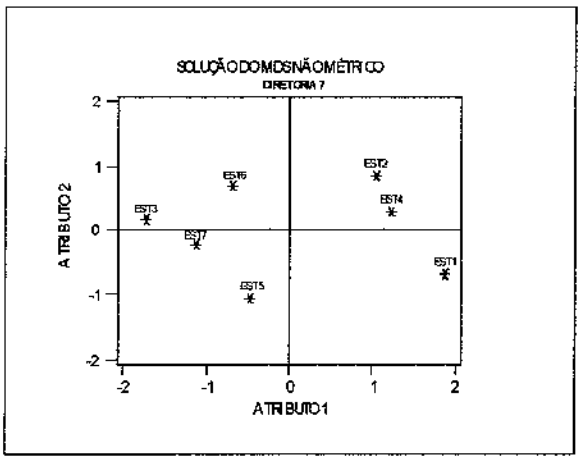


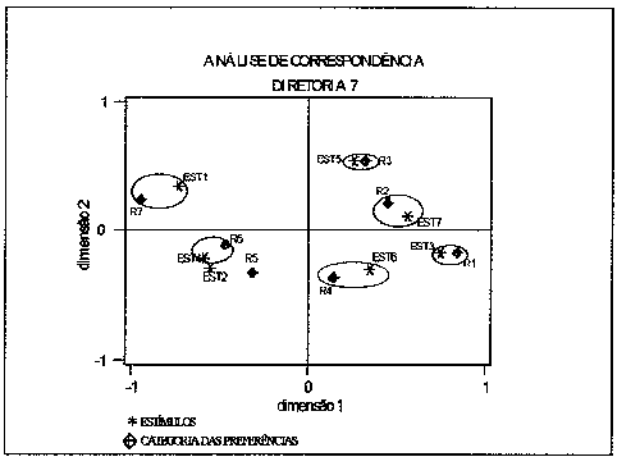
STRESS = 0.02223



% INÉRCIA = 85.35 %

Diretoria 7

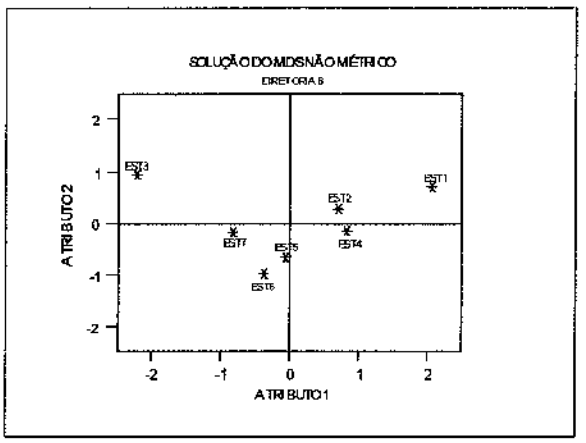


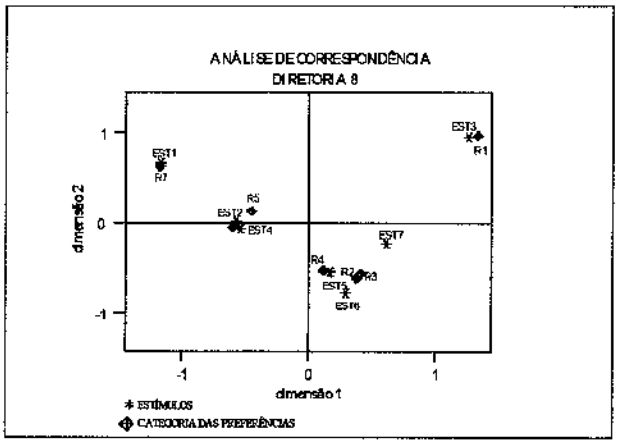


☞ **STRESS = 0.03417**

☞ **% INÉRCIA = 75.03 %**

Diretoria 8

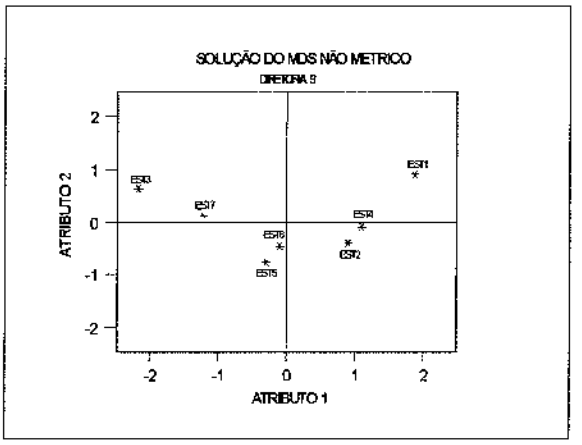


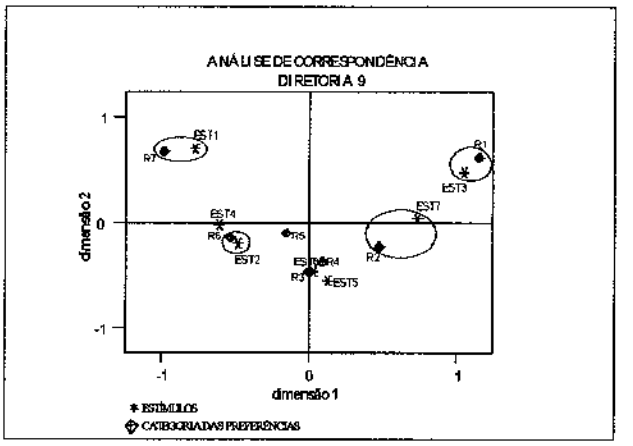


☞ **STRESS = 0.00757**

☞ **% INÉRCIA = 89.23 %**

Diretoria 9

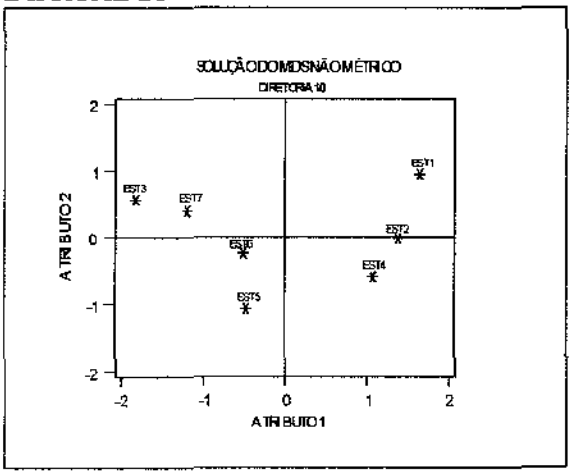




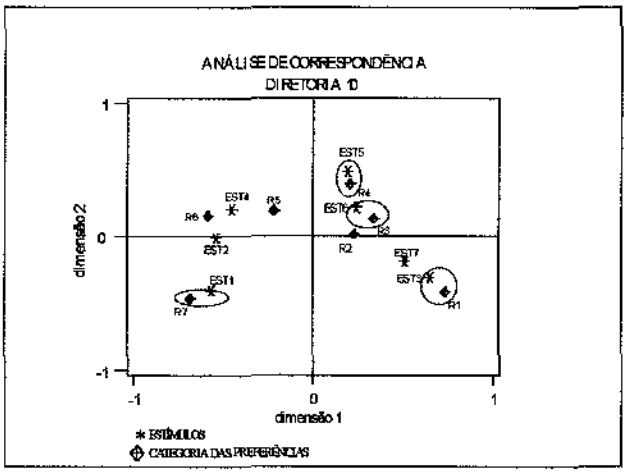
☞ **STRESS = 0.00802**

☞ **% INÉRCIA = 88.20 %**

Diretoria 10

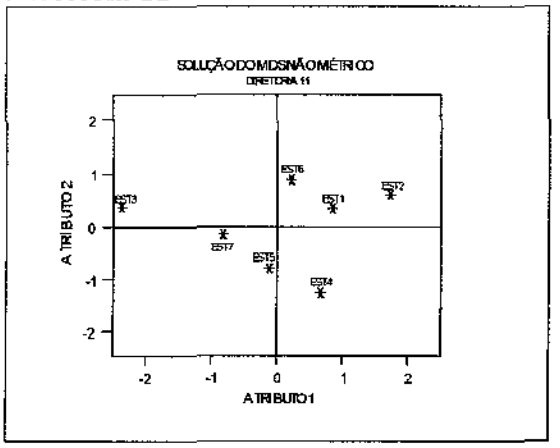


☞ **STRESS = 0.00273**

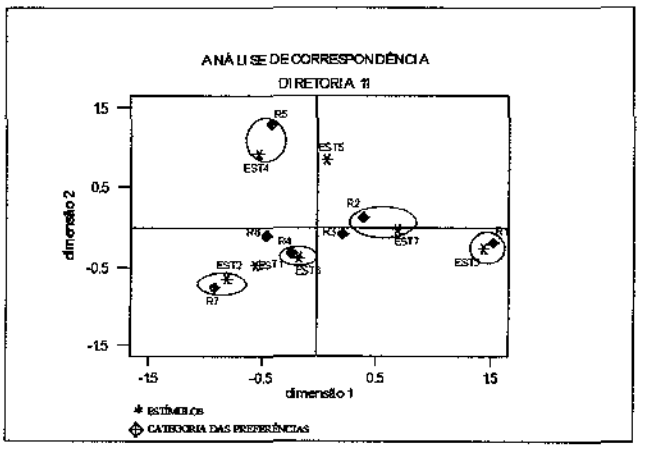


☞ **% INÉRCIA = 86.13 %**

Diretoria 11

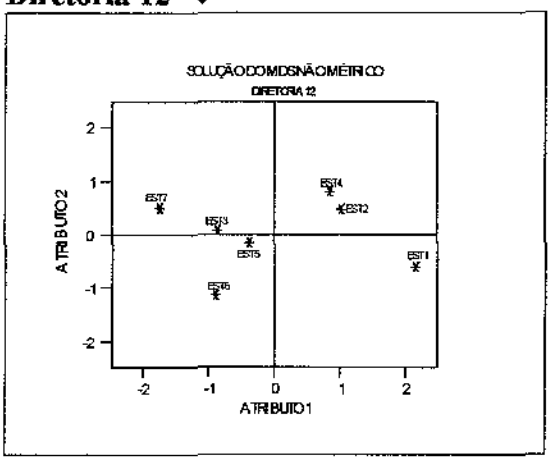


☞ **STRESS = 0.02747**

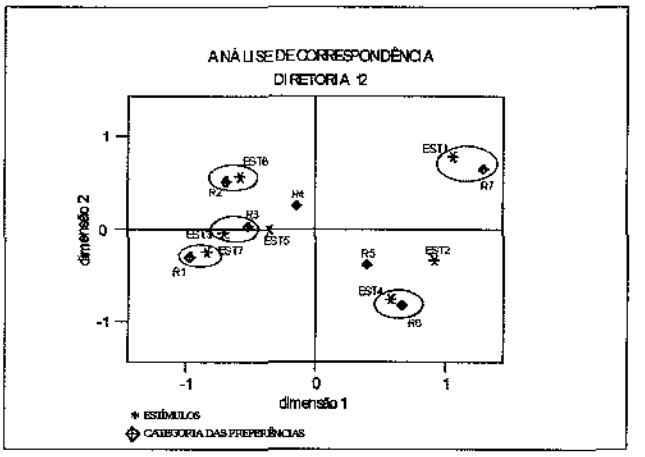


☞ **% INÉRCIA = 64.98 %**

Diretoria 12



☞ **STRESS = 0.01441**



☞ **% INÉRCIA = 69.78 %**

5.5. Resultados via MDS para o Total sem Considerar Estrutura de Grupos.

O desenvolvimento desta seção seguirá os mesmos passos da seção 5.4. A única diferença é que a análise será feita para o total dos 513 indivíduos sem considerar estrutura de grupos.

- Passo 1

Iniciamos a análise trabalhando com a tabela 5.2.

- Passo 2

Neste passo, calculamos a matriz de frequências a partir da tabela 5.2, que forneceu o seguinte resultado.

Tabela 5.6. Freqüências para o total de indivíduos sem considerar estrutura de grupos.

	EST1	EST2	EST3	EST4	EST5	EST6	EST7	n_i	n_i / n
R1	19	10	250	5	32	47	150	513	0.1429
R2	33	29	78	29	112	98	134	513	0.1429
R3	24	48	63	39	138	123	78	513	0.1429
R4	40	69	59	64	105	111	65	513	0.1429
R5	66	113	37	117	76	61	43	513	0.1429
R6	94	113	20	161	38	53	34	513	0.1429
R7	237	131	6	98	12	20	9	513	0.1429
n_j	513	513	513	513	513	513	513	3591	0.1429
n_j / n	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429		

- Passo 3

Agora, a partir da tabela de freqüências, calcularemos a distância Qui-quadrada entre os estímulos. Os resultados são apresentados a seguir:

Matriz 5.2. Matriz de distância Qui-quadrada entre estímulos sem considerar estrutura de grupo.

$$D_{QUI-Q} = \begin{array}{c} \\ EST1 \\ EST2 \\ EST3 \\ EST4 \\ EST5 \\ EST6 \\ EST7 \end{array} \begin{pmatrix} 0 & & & & & & \\ 0.4074 & 0 & & & & & \\ 3.1097 & 2.4032 & 0 & & & & \\ 0.7292 & 0.0941 & 2.5998 & 0 & & & \\ 2.0605 & 1.0048 & 1.5503 & 1.1515 & 0 & & \\ 1.8254 & 0.8547 & 1.3235 & 0.9752 & 0.0318 & 0 & \\ 2.3138 & 1.5307 & 0.3626 & 1.6778 & 0.5510 & 0.4481 & 0 \end{pmatrix}_{7x7}$$

- **Passo 4**

A solução obtida através do MDS não métrico para o total de indivíduos, sem considerar estrutura de grupos, é dado por:

Tabela 5.7. Solução do MDS não métrico do total de 513 indivíduos

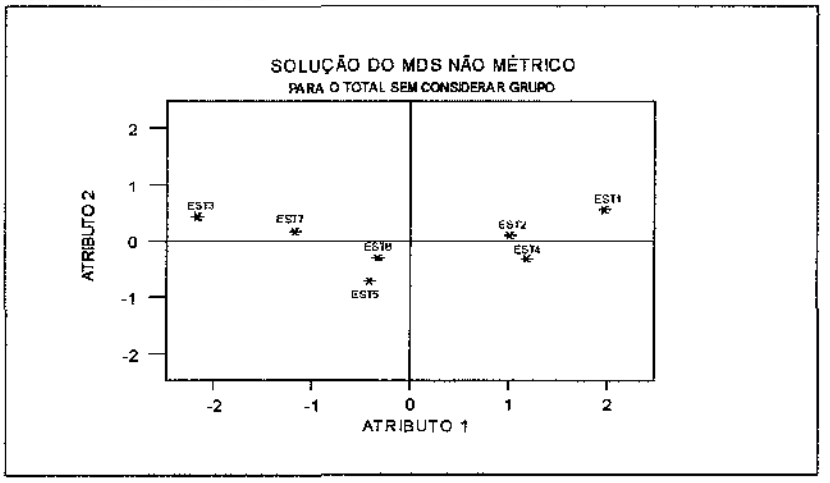
ESTÍMULO	dimensão 1	dimensão 2
EST1	1.96312	0.56166
EST2	1.00406	0.12184
EST3	-2.18915	0.46838
EST4	1.16638	-0.32750
EST5	-0.42847	-0.71627
EST6	-0.33515	-0.29818
EST7	-1.18080	0.19007

A medida do bondade de ajuste, chamada de STRESS foi igual a: 0.00948, a qual está indicando que o ajuste da configuração dos estímulos foi muito bom.

- **Passo 5**

A seguir mostramos um gráfico para poder ter uma idéia do comportamento dos estímulos em um plano bidimensional, usando a informação conjunta dos 513 indivíduos.

Gráfico 5.6. Comportamento dos Estímulos sem considerar estrutura de grupos do Total 513 indivíduos



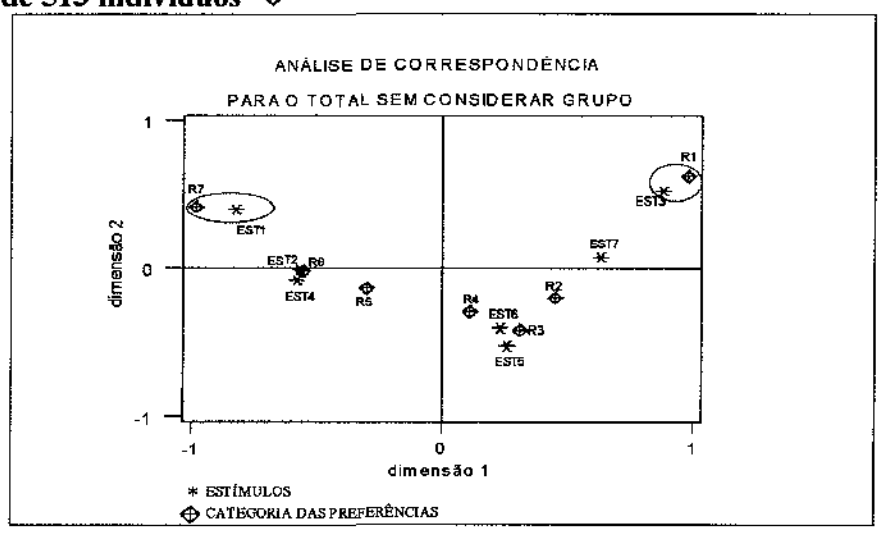
STRESS = 0.00948

Como se poderá observar no gráfico, os estímulos 3 e 7 estão em contra-posição aos estímulos 1 e 4; além disso, os estímulos 6 e 5 estão novamente bastante próximos, indicando que esses estímulos têm algo em comum.

- Passo 6

Adicionalmente, foi usada a Análise de Correspondência para poder comparar os resultados da solução obtida pelo MDS não métrico com a solução dada pela Análise de Correspondência. O resultado foi o seguinte:

Gráfico 5.7. Comportamento conjunto dos estímulos e as frequências das preferências para o total de 513 indivíduos



% DE INÉRCIA = 91.42%

Observa-se que a porcentagem de inércia explicada pelas duas dimensões foi de 91.42%, a qual estaria indicando um bom ajuste dado pelas duas dimensões. Como se poderá observar no gráfico, o estímulo 3 está junto a r1 o qual está nos indicando que o estímulo 3 foi o mais preferido pelos indivíduos, e o estímulo 1 o menos preferido.

Comparando o gráfico dado pelo MDS e o gráfico dado pela Análise de Correspondência, percebe-se que um deles é reflexo do outro, isto é, os estímulos têm quase o mesmo comportamento em ambas as análises.

5.6. Resultados via MDS Considerando Estrutura de Grupos.

Como visto no capítulo III, nesta seção trabalharemos com os 513 indivíduos, porém considerando que eles pertencem a grupos diferentes. Isto é, a técnica de MDS sobre estrutura de grupos nos permite encontrar uma configuração de pontos comums dos estímulos para o total (513 indivíduos), e também nos permite encontrar uma configuração de pontos dos estímulos para cada grupo (diretoria). A idéia desta análise é poder comparar os critérios das seções 5.4 e 5.5 com o critério desta seção 5.6.

Apenas lembrando, o modelo é o seguinte (ver equação 3.3.2 no capítulo III):

$$X_S = XW_S \tag{5.1}$$

onde:

X_s : Configuração dos estímulos para cada grupo.

X : Configuração dos estímulos comums para o total.

W_s : Indica os pesos associados a cada grupo.

para todo s=1,2,..,12.

Esta seção 5.6 é composta de duas partes. A primeira parte é formada pela solução da configuração dos estímulos comums para o total (X) com seus respectivos pesos (W_s). Na segunda parte estimaremos a configuração dos estímulos de cada grupo a partir da solução da primeira parte.

5.6.1. Primeiro Critério.

A solução do MDS sobre estrutura de grupos a matriz comum X é dada por:

Tabela 5.8. Solução do MDS sobre estrutura de grupos 513 indivíduos

ESTÍMULO	dimensão 1	dimensão 2
EST1	0.98854	1.20943
EST2	1.10307	0.43797
EST3	-1.36786	1.29185
EST4	1.13100	-0.72014
EST5	-0.45422	-1.65857
EST6	-0.21614	0.07331
EST7	-1.18439	-0.63384

A medida de bondade do ajuste para este caso foi: S-STRESS=0.06432

Os pesos respectivos para cada grupo foram:

$$W_1 = \begin{pmatrix} 1.27 & 0 \\ 0 & 0.62 \end{pmatrix}_{2x2} \qquad W_2 = \begin{pmatrix} 1.34 & 0 \\ 0 & 0.44 \end{pmatrix}_{2x2} \qquad W_3 = \begin{pmatrix} 1.02 & 0 \\ 0 & 0.98 \end{pmatrix}_{2x2}$$

$$W_4 = \begin{pmatrix} 1.28 & 0 \\ 0 & 0.61 \end{pmatrix}_{2x2} \qquad W_5 = \begin{pmatrix} 1.28 & 0 \\ 0 & 0.60 \end{pmatrix}_{2x2} \qquad W_6 = \begin{pmatrix} 1.13 & 0 \\ 0 & 0.85 \end{pmatrix}_{2x2}$$

$$W_7 = \begin{pmatrix} 1.23 & 0 \\ 0 & 0.69 \end{pmatrix}_{2x2} \qquad W_8 = \begin{pmatrix} 1.00 & 0 \\ 0 & 1.00 \end{pmatrix}_{2x2} \qquad W_9 = \begin{pmatrix} 1.27 & 0 \\ 0 & 0.62 \end{pmatrix}_{2x2}$$

$$W_{10} = \begin{pmatrix} 1.28 & 0 \\ 0 & 0.60 \end{pmatrix}_{2x2} \qquad W_{11} = \begin{pmatrix} 1.09 & 0 \\ 0 & 0.90 \end{pmatrix}_{2x2} \qquad W_{12} = \begin{pmatrix} 1.10 & 0 \\ 0 & 0.88 \end{pmatrix}_{2x2}$$

5.6.2. Segundo Critério.

A partir da equação (5.1) podemos estimar a configuração para cada diretoria, pois basta multiplicar a configuração comum dos estímulos por seus respectivos pesos. Assim, a configuração dos estímulos para a diretoria 1 é dado por:

Tabela 5.9. Solução do MDS sobre estrutura de grupos para diretoria 1

ESTÍMULO	dimensão 1	dimensão 2
EST1	1.2554	0.7498
EST2	1.4009	0.2715
EST3	-1.7372	0.8009
EST4	1.4364	-0.4465
EST5	-0.5769	-1.0283
EST6	-0.2744	0.0455
EST7	-1.5042	-0.3930

Usando o mesmo procedimento, estimaram-se as configurações para as 12 diretorias. A seguir mostramos a configuração dos estímulos para a diretoria 12:

Tabela 5.10. Solução do MDS sobre estrutura de grupos para diretoria 12.

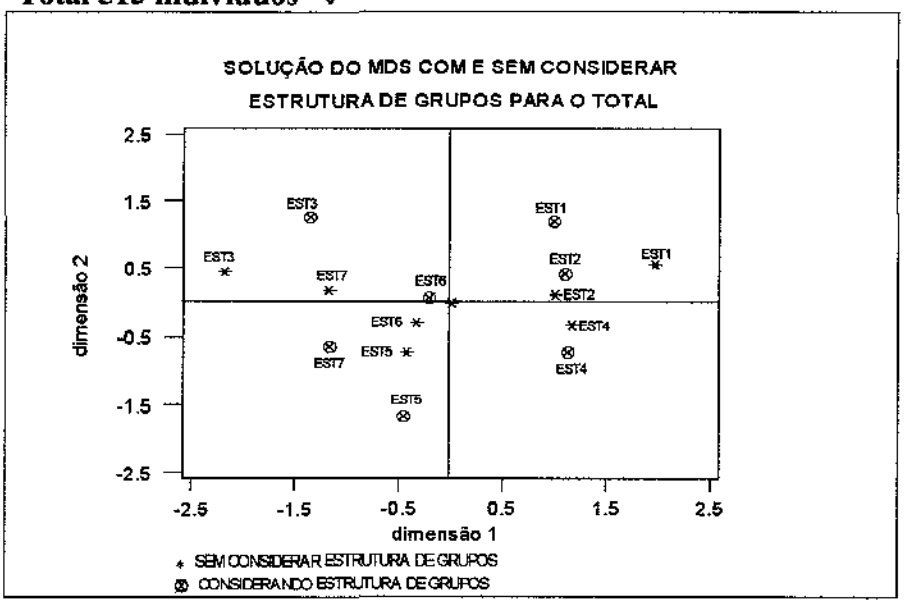
Diretoria 12 sobre estrutura de grupos

ESTÍMULO	dimensão 1	dimensão 2
EST1	1.0874	1.0643
EST2	1.2133	0.3854
EST3	-1.5046	1.1368
EST4	1.2441	-0.6337
EST5	-0.4996	-1.4595
EST6	-0.2377	0.0645
EST7	-1.3028	-0.55784

Nesta parte, nosso interesse é comparar as seções 5.4, 5.5 e 5.6, para observar se realmente existe uma diferença acentuada entre grupos, e também poder perceber como essas diferenças estão sendo refletidas. A seguir, mostramos os gráficos conjuntos das seções 5.5 e 5.6.1, e também as seções 5.4 com 5.6.2. Isto é, em um só gráfico apresentaremos ambos os critérios para podermos observar o que está acontecendo com o comportamento dos estímulos através dos resultados obtidos nas seções anteriores.

Gráfico 5.8. Gráfico comparativo das soluções considerando e não consideraando estrutura de grupos para o total de indivíduos.

Total 513 indivíduos



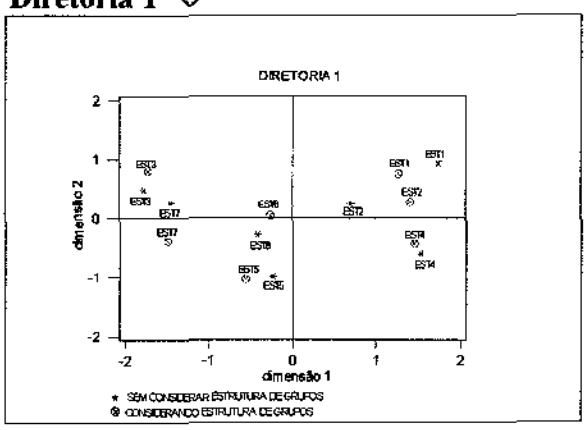
STRESS = 0.00948 **S-STRESS = 0.06432**

Observa-se no gráfico 5.8, que o comportamento dos estímulos é quase o mesmo. A mudança mais acentuada é notada no estímulo 1. Por exemplo, nota-se que em ambos os casos a dimensão 1 está medindo a ordem de preferências dos 513 inívíduos pelos estímulos. Estes aspectos serão detalhados na conclusão final.

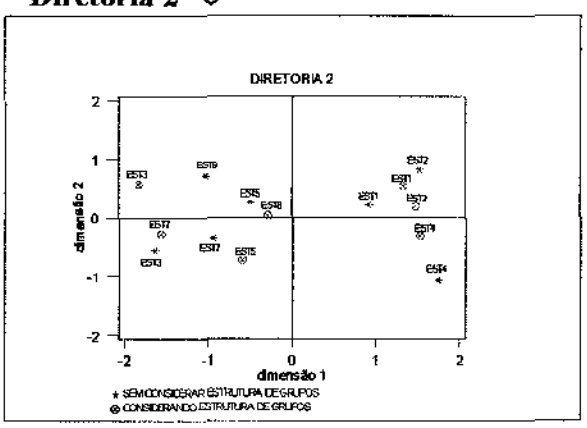
A seguir, mostraremos uma síntese dos gráficos conjuntos das soluções do MDS considerando e não considerando estrutura de grupos para cada diretoria. Basicamente, estes gráficos foram construídos para podermos observar o comportamento dos estímulos através das soluções obtidas das seções 5.4 e 5.6.2.

Gráfico 5.9. Sínteses do comportamento dos estímulos considerando e não considerando estrutura de grupos por diretorias.

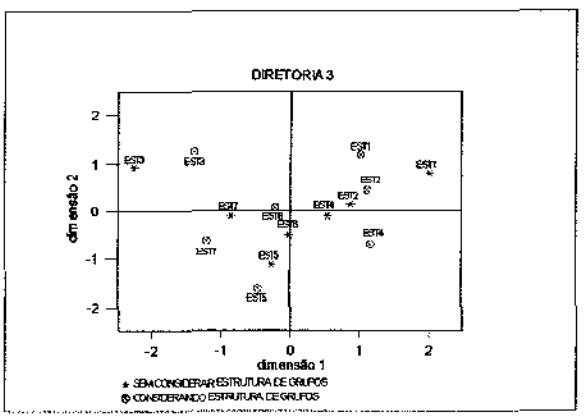
Diretoria 1



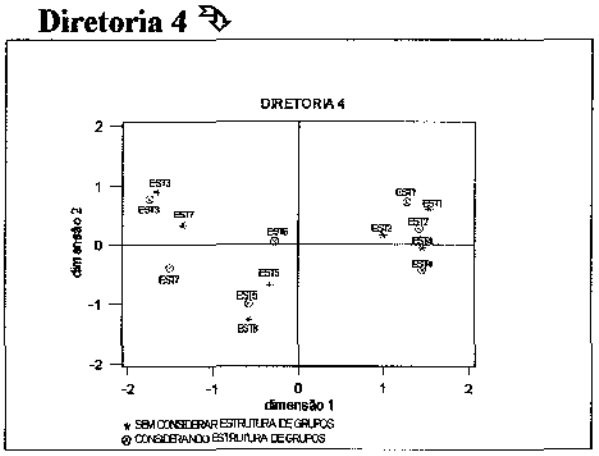
Diretoria 2



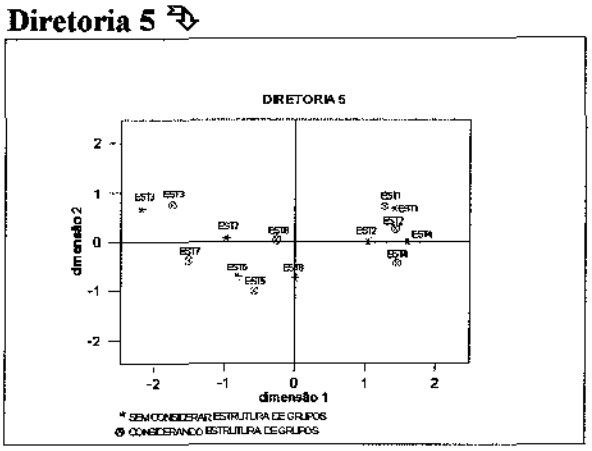
Diretoria 3



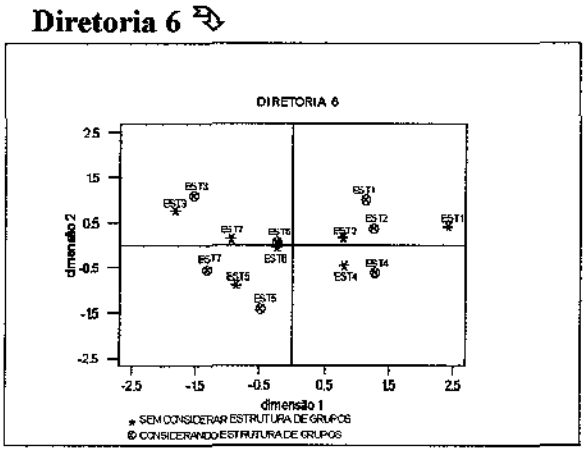
Diretoria 4



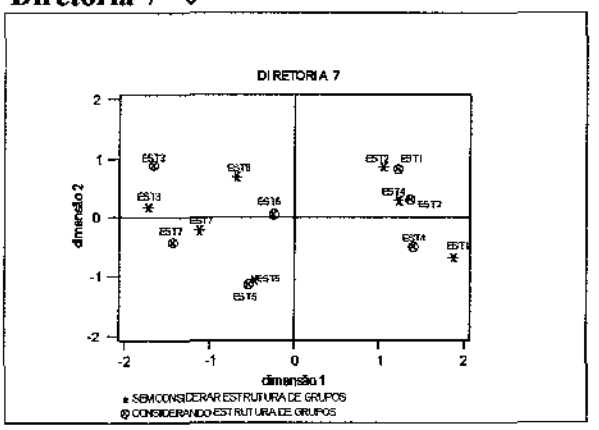
Diretoria 5



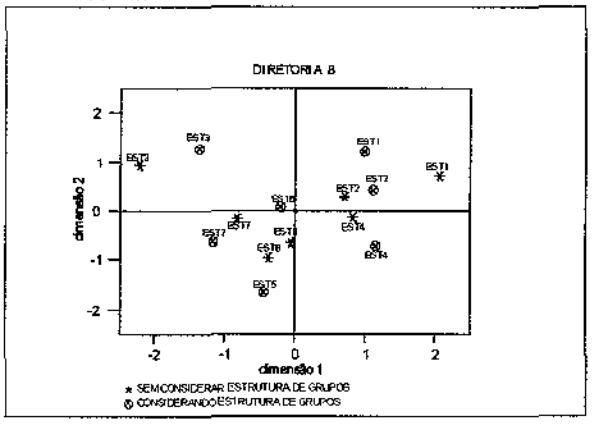
Diretoria 6



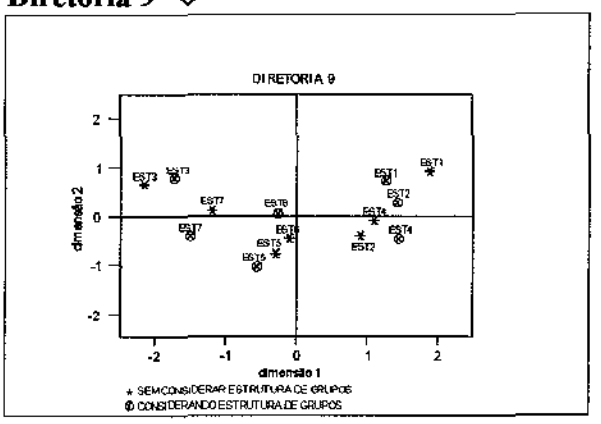
Diretoria 7



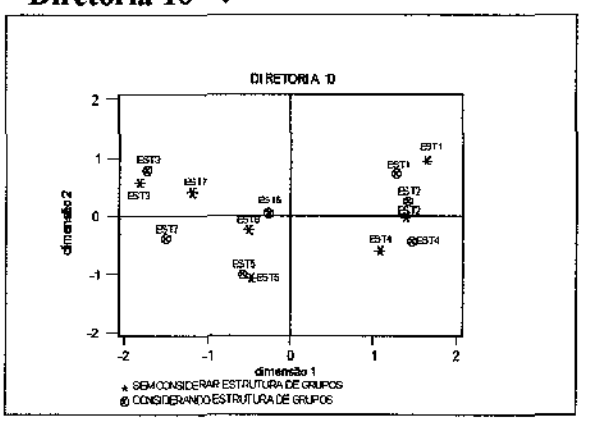
Diretoria 8



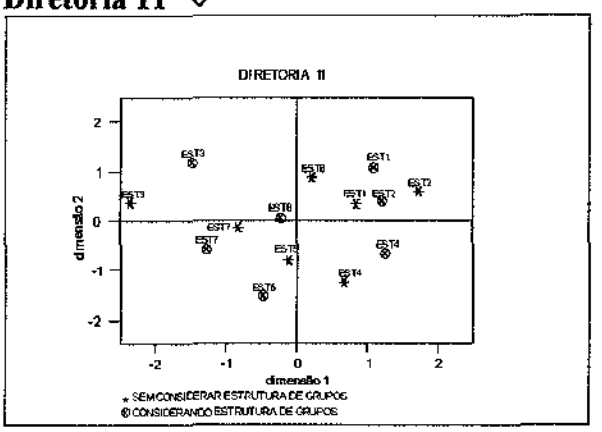
Diretoria 9



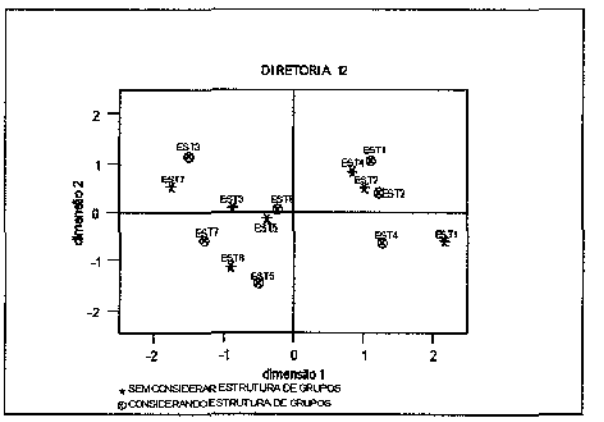
Diretoria 10



Diretoria 11



Diretoria 12



5.7. Análise dos Resultados Obtidos Através do MDS.

A estratégia para solução do problema de estudar as preferências, via MDS, foi feita a partir de três pontos de vista diferentes:

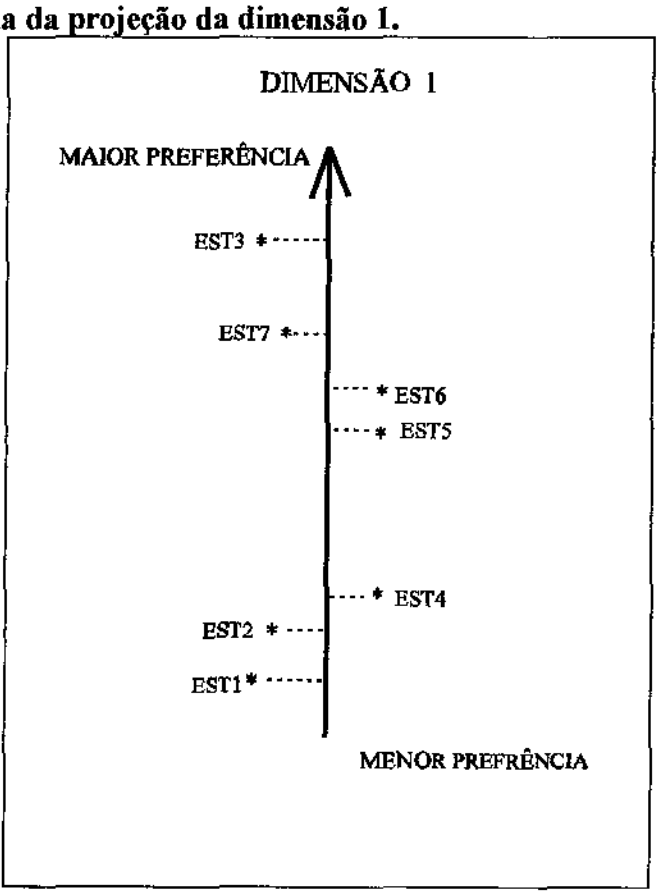
1 - enfocando indivíduos dentro de cada diretoria,

2 - para o conjunto total, onde analisamos todos os indivíduos,

3 - considerando a estrutura dos grupos (diretorias).

Claramente, as matrizes de distâncias, em cada caso, foram obtidas de forma distinta e os resultados em cada caso poderiam ser distintos. Na seção 5.4, através da síntese dos gráficos (ver gráfico 5.5) observa-se que a configuração dos estímulos via MDS para as diretorias apresentam um padrão que se reproduz com pequenas variações: é com regularidade que os estímulos apresenta uma ordem de projeção na primeira dimensão, dada pelo estímulo 3, seguida do estímulo7, depois pelo par de estímulos 5 e 6, e, em oposição, aparecem os estímulos 1, 2 e 4. A seguir mostramos um diagrama, dando uma idéia melhor do que está acontecendo.

Gráfico 5.10. Esquema da projeção da dimensão 1.



Nota-se que os estímulos mais preferidos são 3 e 7, isto é, os estímulos 3 e 7 são os mais importantes, seguidos pelos estímulos 5 e 6. Às vezes, os estímulos 5 e 6 mudam a posição, isto é, pode ser 6 e 5, (veja que eles sempre ficam próximos; isto está indicando algo em comum entre eles). Os estímulos menos preferidos foram os estímulos 4, 2 e 1. Às vezes, há mudança de posição podendo ser 2, 4 e 1, mas sempre estes 3 últimos estímulos figuram como os menos preferidos.

Além disso, a Análise de Correspondência exibe o mesmo padrão para os estímulos, uma correspondência com as prefrências. Também percebe-se sistematicamente associados às maiores preferências os estímulos 3 e 7 e às menores preferências os estímulos 1,2, e 4.

Pela combinação de ambos os resultados, podemos concluir que há uma ordem de preferência ao longo da dimensão 1, e que se reproduz através das diretorias. Este resultado permite antever os demais resultados, pois a análise do total e a análise cosiderando estrutura de grupo só acrescentariam informação no caso dos grupos apresentarem padrões distintos na ordem de preferências. Neste sentido, o critério da formação das diretorias não se revela como fonte de variação no que diz respeito à avaliação institucional.

Os resultados das seções 5.5 e 5.6 foram mantidos ao longo de toda a completitude do roteiro de análise, uma vez que são informativos caso haja diferença entre grupos. Revelam-se praticamente os mesmos, como mostra a síntese gráfica comparativa. A seguir nota-se que as configurações dos estímulos obtidos, sem considerar e considerando estrutura de grupos, são semelhantes, o que é esperado. Caso isto não ocorresse, as diferenças das matrizes de pesos obtidos na seção 5.6.1 ajudariam a entender as deferenças entre grupos (diretorias), ainda que as configurações pelos dois métodos fossem semelhantes.

Os resultados relativos ao total apresentado na seção 5.5. são esperados, dado que não foi notada diferença entre grupos. Ressaltamos que a Análise de Correspondência visa revelar as associações entre preferências e estímulos, bem como, através das inércias, complementar a informação dada pelo STRESS, uma vez que as duas dimensões foram, ao longo das análises, responsáveis pela variação entre os estímulos.

5.8. Resultados dos Modelos de Preferência.

Nesta segunda parte, dado que já encontramos uma configuração de pontos para os estímulos, estimaremos os pontos de preferência máxima para cada indivíduo, denominados Pontos Ideais, assim como foi mencionado no capítulo IV. Como vimos, existem 4 modelos de preferência para estimar os pontos ideais de cada indivíduo. Nesta aplicação, foi usado o modelo Ponto Ideal Ponderado, pelas seguintes razões:

a.- O primeiro modelo de preferência Vetor (linear) não foi usado pelo fato das respostas dos indivíduos não estarem linearmente relacionadas com os atributos.

b.- O segundo modelo de preferência Ponto Ideal Simples (não linear) também foi descartado, porque ele supõe que os atributos têm o mesmo peso. Com a Análise de Correspondência foi descartada essa suposição, isto é, mediante a porcentagem de inércia explicada pelas dimensões da solução da Análise de Correspondência, verificamos que efetivamente cada atributo (dimensão) tinha porcentagem de explicação diferente.

c.- O modelo Ponto Ideal Geral (não linear) também foi descartado, porque os dois atributos encontrados estão explicando coisas diferentes. Além disso, nos gráficos feitos anteriormente, notou-se que não existe interação entre os atributos.

Então, o modelo Ponto Ideal Ponderado (não linear) foi considerado para esta aplicação. A análise, nesta parte, foi feita separadamente para cada grupo, pois já vimos na primeira parte que o comportamento dos estímulos é quase o mesmo, considerando ou não considerando estrutura de grupos.

5.8.1. Estimação do Modelo Ponto Ideal Ponderado dos indivíduos para cada diretoria.

DIRETORIA 1

Como sabemos, para poder estimar o Ponto Ideal de cada indivíduo, é preciso estimar um modelo para cada indivíduo. Exemplificaremos estimando apenas para o primeiro, segundo e último indivíduo.

Indivíduo 1

- Da tabela 5.3 vamos extrair a resposta do indivíduo 1, dada por:

$$\delta_{11} = (7 \quad 1 \quad 2 \quad 5 \quad 6 \quad 4 \quad 3)^t$$

- Tomamos a solução do MDS para a diretoria 1 dada pela tabela 5.5 que é a seguinte:

Diretoria 1

ESTÍMULO	ATRIBUTO 1	ATRIBUTO 2
EST1	1.73514	0.92212
EST2	0.68235	0.23358
EST3	-1.79318	0.46527
EST4	1.52092	-0.60518
EST5	-0.25305	-0.98792
EST6	-0.43274	-0.28762
EST7	-1.45943	0.25976

- Vimos no capítulo IV, que depois de reparametrizado, o modelo Ponto Ideal Ponderado é expresso da seguinte maneira:

$$\delta_{is} = w_s \sum_{k=1}^{K} x_{ik}^2 + \sum_{k=1}^{K} \beta_{ks} x_{ik} + \varepsilon_s$$

onde os parâmetros serão estimados usando Regressão Monotônica (ver apêndice E), já que os dados foram medidos em uma escala Ordinal. Para entender melhor esta parte é recomendável ver os passos a seguir, proposto no capítulo IV na parte de estimação não métrica.

Na parte de estimação do modelo, foram necessárias 30 iterações para estimar o modelo final, dado por:

$$\hat{\delta}_{11} = 1.88606 + 0.50749\text{atrib1} - 0.80057\text{atrib2} + 0.33210\text{atrib1}^2 + 4.21468\text{atrib2}^2$$

A partir deste modelo, podemos estimar o ponto ideal ponderado para o indivíduo 1, da seguinte maneira:

$$\hat{\pi}_{11} = \frac{0.50749}{(-2)(0.33210)} = -0.76406 \qquad \hat{\pi}_{12} = \frac{-0.80057}{(-2)(4.21468)} = -0.09497$$

A seguir, mostramos as contribuições dos atributos para o modelo,

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.50749	38.78%
Atributo 2	-0.80057	61.22%

Pode-se observar que o atributo 2 contribuiu mais para a estimação do modelo, isto é, para estimar o ponto ideal ponderado do indivíduo 1, o atributo 2 teve maior peso.

Analogamente, podemos estimar o ponto ideal do indivíduo 2.

Indivíduo 2

- Da tabela 5.3 vamos extrair a resposta do indivíduo 2, dado por:

$$\delta_{21} = (7 \quad 5 \quad 4 \quad 6 \quad 1 \quad 3 \quad 2)^t$$

- Agora tomamos a solução do MDS da diretoria 1, dado pela tabela 5.5 que é a seguinte:

Diretoria 1 ➪

ESTÍMULO	ATRIBUTO 1	ATRIBUTO 2
EST1	1.73514	0.92212
EST2	0.68235	0.23358
EST3	-1.79318	0.46527
EST4	1.52092	-0.60518
EST5	-0.25305	-0.98792
EST6	-0.43274	-0.28762
EST7	-1.45943	0.25976

Na parte de estimação do modelo, foram necessárias 24 iterações para estimar o modelo final, dado por:

$$\hat{\delta}_{21} = 3.75242 + 1.44034\text{atrib1} + 0.75864\text{atrib2} + 0.53475\text{atrib1}^2 - 1.66648\text{atrib2}^2$$

A partir deste modelo podemos estimar o ponto ideal ponderado para o indivíduo 2, da seguinte maneira:

$$\hat{\pi}_{21} = \frac{1.44034}{(-2)(0.53475)} = -1.34674 \qquad \hat{\pi}_{22} = \frac{0.75864}{(-2)(-1.66648)} = 0.22762$$

Mostramos a seguir as contribuições dos atributos para o modelo, as quais foram:

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.44034	65.50%
Atributo 2	0.75864	34.49%

Pode-se observar que o atributo 1 contribuiu mais para a estimação do modelo, isto é, para estimar o ponto ideal ponderado do indivíduo 2, o atributo 1 teve maior peso. Assim sucessivamente, estima-se os modelos para todos os indivíduos. Para o último indivíduo da diretoria 1 temos:

Indivíduo 49

- Da tabela 5.5 vamos extrair a resposta do indivíduo 49, que é a seguinte:

$$\delta_{491} = (7 \quad 4 \quad 6 \quad 5 \quad 2 \quad 3 \quad 1)^t$$

- Agora tomamos a solução do MDS da diretoria 1, dado por:

Diretoria 1

ESTÍMULO	ATRIBUTO 1	ATRIBUTO 2
EST1	1.73514	0.92212
EST2	0.68235	0.23358
EST3	-1.79318	0.46527
EST4	1.52092	-0.60518
EST5	-0.25305	-0.98792
EST6	-0.43274	-0.28762
EST7	-1.45943	0.25976

Na parte de estimação do modelo, foram necessárias 13 iterações para estimar o modelo final, dado por:

$$\hat{\delta}_{491} = 2.86723 + 0.72844\text{atrib1} + 2.18735\text{atrib2} + 0.19766\text{atrib1}^2 + 2.17266\text{atrib2}^2$$

A partir deste modelo, podemos estimar o ponto ideal ponderado para o indivíduo 49 da seguinte maneira:

$$\hat{\pi}_{491} = \frac{0.72844}{(-2)(0.19766)} = -1.84267 \qquad \hat{\pi}_{492} = \frac{2.18735}{(-2)(2.17266)} = -0.50337$$

Mostramos, a seguir, as contribuições dos atributos para o modelo:

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.72844	24.98%
Atributo 2	2.18735	75.02%

Pode-se observar que o atributo 2 contribuiu mais para a estimação do modelo, isto é, para estimar o ponto ideal ponderado do indivíduo 49, o atributo 2 teve maior peso.

O SAS calcula direto os pontos ideais ponderados para cada indivíduo (ver apêndice G).

5.8.2. Modelo Ponto Ideal Ponderado para cada diretoria.

Como se pôde observar até agora, estimamos um modelo diferente para cada indivíduo. Agora que temos um modelo para cada indivíduo, consequentemente a idéia é encontrar um modelo para cada diretoria. Isto será feito usando os modelos estimados dos indivíduos. Por exemplo, na diretoria 1 temos 49 indivíduos, consequentemente temos 49 modelos estimados. Então, para estimar um só modelo representando a diretoria 1, tomamos as médias de cada parâmetro estimado. Assim, o modelo estimado para a diretoria 1 foi o seguinte:

$$\hat{\delta}_{Dir1} = 3.831 + 0.9101\text{atrib1} - 0.171\text{atrib2} + 0.085\text{atrib1}^2 + 0.085\text{atrib2}^2$$

A partir deste modelo podemos estimar o ponto ideal ponderado para a diretoria 1 da seguinte maneira:

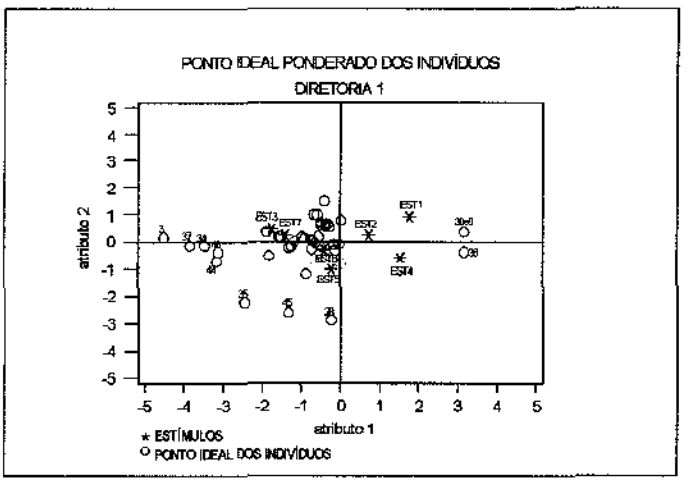
$$\hat{\pi}_{Dir11} = \frac{0.9101}{(-2)(0.085)} = -5.35350 \qquad \hat{\pi}_{Dir12} = \frac{0.171}{(-2)(0.085)} = 1.00588$$

A seguir mostramos a contribuição de cada atributo para o ajuste do modelo.

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.9101	84.18%
Atributo 2	-0.1710	15.82%

O resultado final para a diretoria 1 foi o seguinte:

Gráfico 5.11. Resultado final dos pontos ideais da diretoria 1



☞ **STRESS = 0.00390** ☞ **% ATRIB. 1 = 84.18%**

Menciona-se que no gráfico encontram-se os pontos ideais de 45 indivíduos. 4 indivíduos não foram levados em conta, uma vez que os pontos ideais deles eram extremamente grandes, caindo fora do gráfico 5.11. O mesmo aconteceu nas demais diretorias, isto é, em média de 3 indivíduos por diretorias não foram levados em conta, com exceção das diretorias 11 e 12.

Analogamente, foram estimados os restantes dos modelos, tanto para os indivíduos como para cada diretoria. A seguir mostramos os modelos estimados para cada diretoria, os

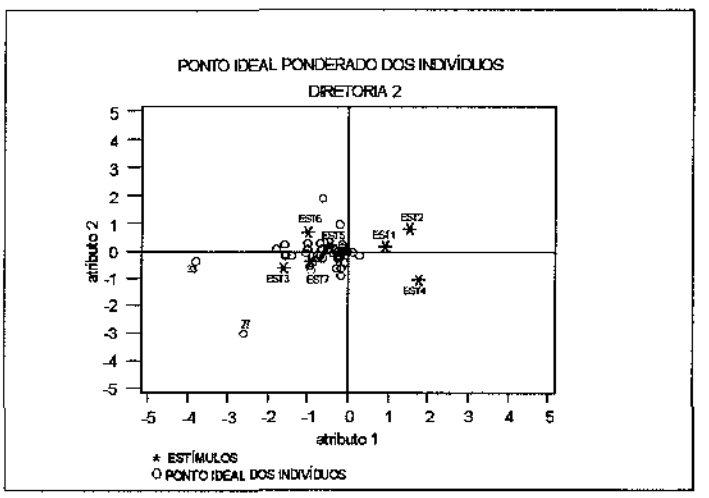
pontos ideais para cada diretoria, assim como a contribuição de cada atributo para o ajuste de modelo.

DIRETORIA 2

$$\hat{\delta}_{Dir2} = 4.370 + 1.0756atrib1 + 0.125atrib2 - 0.201atrib1^2 - 0.122atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.0756	89.59%
Atributo 2	0.125	10.41%

Gráfico 5.12. Resultado final dos pontos ideais da diretoria 2.



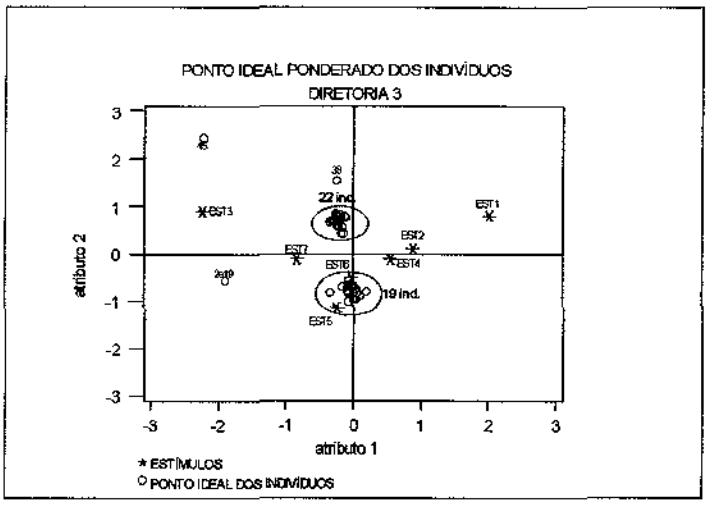
STRESS = 0.01968 **% ATRIB. 1 = 89.59%**

DIRETORIA 3

$$\hat{\delta}_{Dir3} = 4.238 + 1.039atrib1 + 0.34atrib2 - 0.214atrib1^2 + 0.23atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.039	75.34%
Atributo 2	0.34	24.66%

Gráfico 5.13. Resultado final dos pontos ideais da diretoria 3.



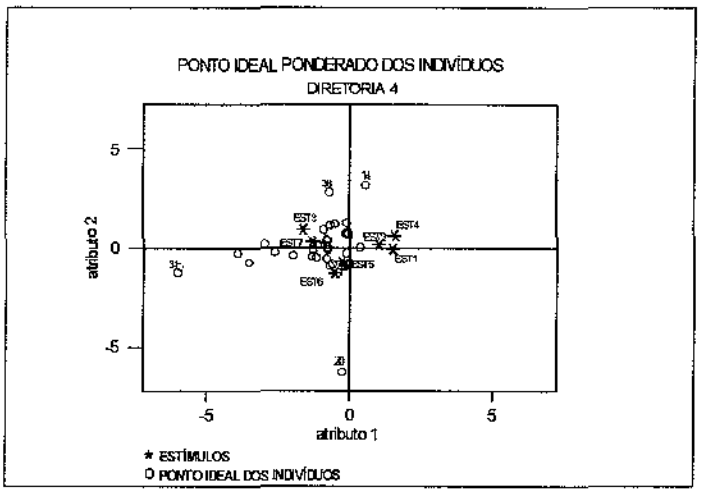
☞ **STRESS = 0.02805** ☞ **% ATRIB. 1 = 75.34%**

DIRETORIA 4 ☜

$$\hat{\delta}_{Dir4} = 3.906 + 1.2054atrib1 + 0.136atrib2 - 0.013atrib1^2 + 0.233atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.2054	89.86%
Atributo 2	0.136	10.14%

Gráfico 5.14. Resultado final dos pontos ideais da diretoria 4.



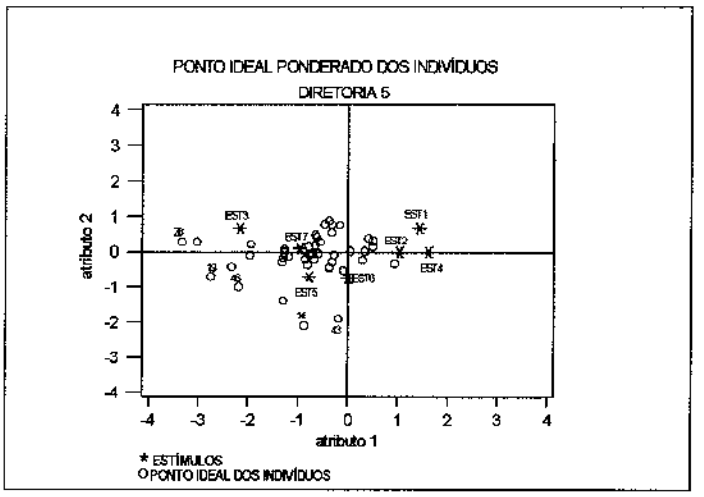
☞ **STRESS = 0.01009** ☞ **% ATRIB. 1 = 89.86%**

DIRETORIA 5

$$\hat{\delta}_{Dir5} = 3.942 + 0.8692atrib1 - 0.322atrib2 + 0.116atrib1^2 - 0.507atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.8692	72.97%
Atributo 2	-0.322	27.03%

Gráfico 5.15. Resultado final dos pontos ideais da diretoria 5.



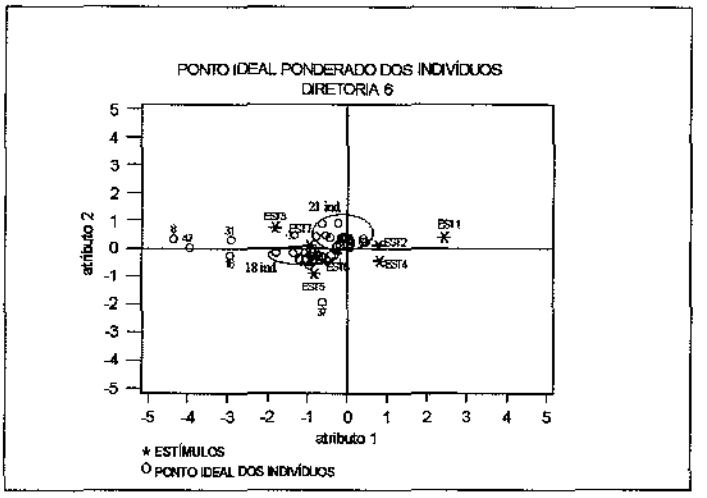
☞ **STRESS = 0.01424** ☞ **% ATRIB. 1 = 72.97%**

DIRETORIA 6

$$\hat{\delta}_{Dir6} = 3.987 + 0.994atrib1 + 0.312atrib2 - 0.077atrib1^2 + 0.566atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.994	76.11%
Atributo 2	0.312	23.89%

Gráfico 5.16. Resultado final dos pontos ideais da diretoria 6.



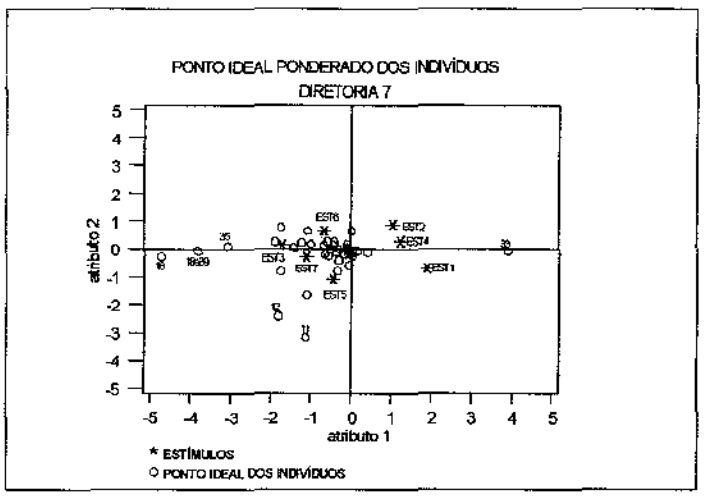
☞ **STRESS = 0.02223** ☞ **% ATRIB. 1 = 76.11%**

DIRETORIA 7 ⇘

$$\hat{\delta}_{Dir7} = 4.057 + 0.95549atrib1 + 0.119atrib2 - 0.049atrib1^2 + 0.047atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.95549	88.92%
Atributo 2	0.119	11.08%

Gráfico 5.17. Resultado final dos pontos ideais da diretoria 7.



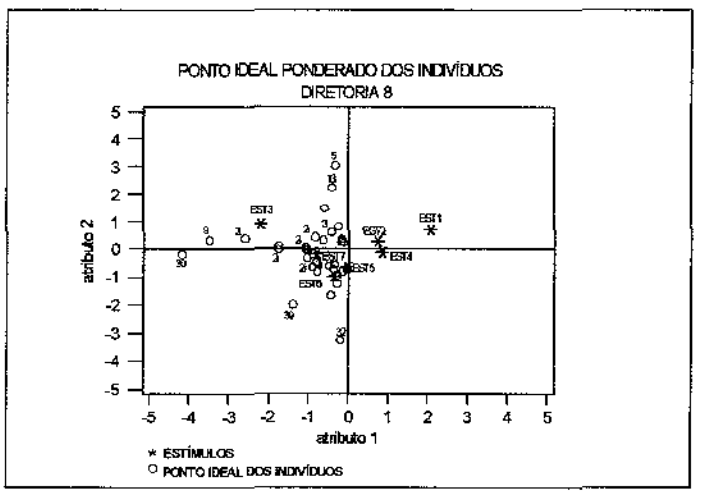
☞ **STRESS = 0.03417** ☞ **% ATRIB. 1 = 88.92%**

DIRETORIA 8

$$\hat{\delta}_{Dir8} = 4.229 + 1.2094atrib1 + 0.955atrib2 - 0.306atrib1^2 + 0.635atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.2094	55.88%
Atributo 2	0.955	44.12%

Gráfico 5.18. Resultado final dos pontos ideais da diretoria 8.



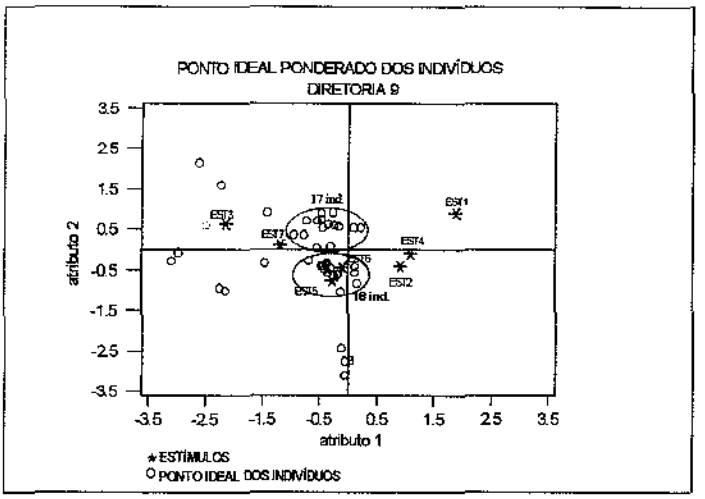
STRESS = 0.00757 **% ATRIB. 1 = 55.88%**

DIRETORIA 9

$$\hat{\delta}_{Dir9} = 4.098 + 0.956atrib1 - 0.025atrib2 - 0.004atrib1^2 - 0.286atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.956	97.45%
Atributo 2	-0.025	2.55%

Gráfico 5.19. Resultado final dos pontos ideais da diretoria 9.



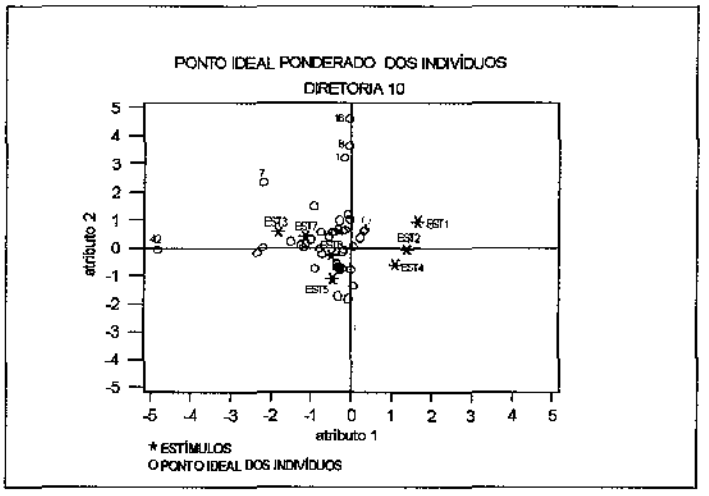
☞ **STRESS = 0.00802** ☞ **% ATRIB. 1 = 97.45%**

DIRETORIA 10 ➪

$$\hat{Y}_{Dir10} = 3.830 + 0.7949atrib1 - 0.230atrib2 + 0.078atrib1^2 + 0.112atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	0.7949	77.56%
Atributo 2	-0.230	22.44%

Gráfico 5.20. Resultado final dos pontos ideais da diretoria 10.



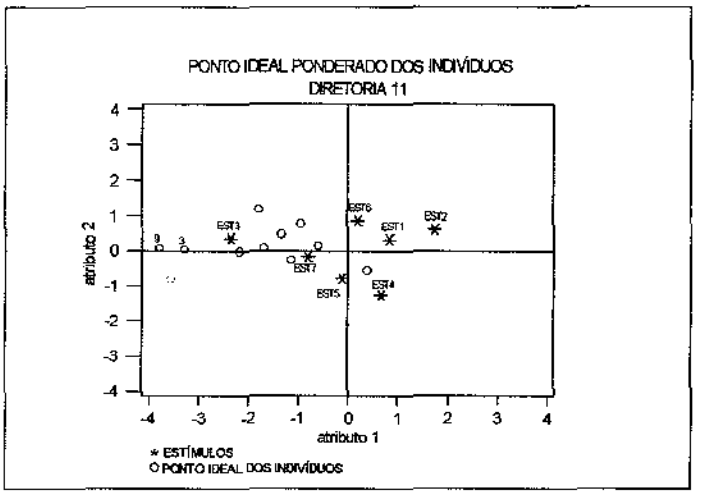
☞ **STRESS = 0.00273** ☞ **% ATRIB. 1 = 77.56%**

DIRETORIA 11

$$\hat{Y}_{Dir11} = 3.553 + 1.288atrib1 + 0.054atrib2 + 0.206atrib1^2 + 0.275atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.288	95.98%
Atributo 2	0.054	4.02%

Gráfico 5.21. Resultado final dos pontos ideais da diretoria 11.



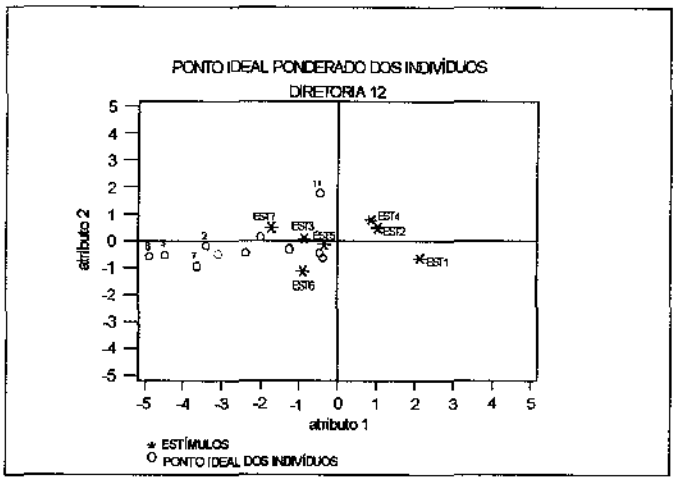
☞ **STRESS = 0.02747** ☞ **% ATRIB. 1 = 95.98%**

DIRETORIA 12

$$\hat{Y}_{Dir12} = 3.983 + 1.2365atrib1 + 0.349atrib2 - 0.053atrib1^2 + 0.248atrib2^2$$

ATRIBUTO	ESTIMATIVA	I_K
Atributo 1	1.2365	77.99%
Atributo 2	0.349	22.01%

Gráfico 5.22. Resultado final dos pontos ideais da diretoria 12.

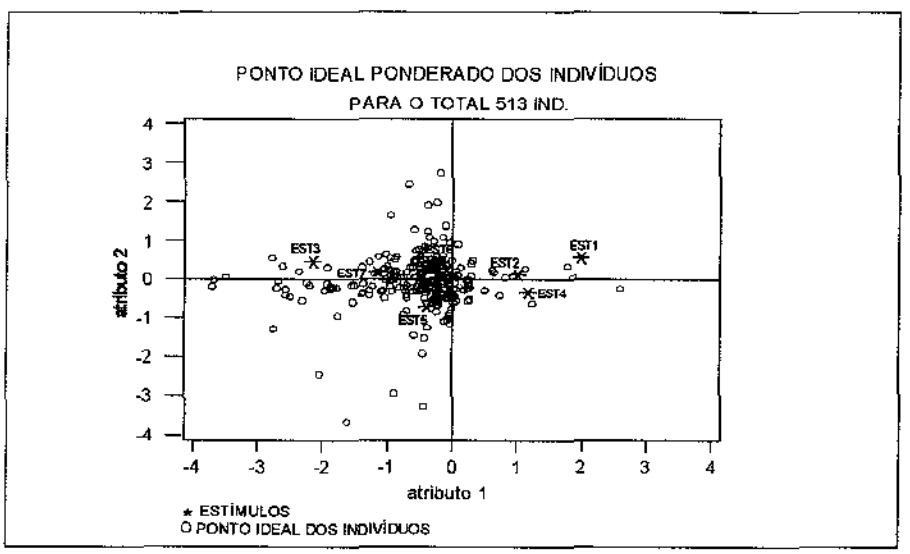


☞ **STRESS = 0.01441** ☞ **% ATRIB. 1 = 77.99%**

Nota-se que em todos os modelos estimados o atributo 1 está contribuindo mais no ajuste. Ou seja, o atributo 1 está dando maior peso na estimação dos pontos ideias. Isto pode ser percebido através dos gráficos 5.11-5.22, pois todos os pontos ideais estão mais ligados ao atributo 1.

Um procedimento análogo foi feito para o total de indivíduos, isto é, estimamos o Ponto Ideal Ponderado dos 513 indivíduos, usando a solução do MDS sem considerar estrutura de grupos. A solução obtida foi a seguinte:

Gráfico 5.24. Resultado final dos pontos ideais para o total dos indivíduos.



☞ **STRESS = 0.00948** ☞ **% ATRIB 1. = 0.06432**

5.9. Análise dos Resultados dos Modelos de Preferência.

Os resultados obtidos nesta parte, possibilita-nos ter uma idéia gráfica do comportamento dos indivíduos frente aos estímulos. A estratégia adotada revela o posicionamento do indivíduo com respeito a suas preferências máximas. Além disso, esta análise permite calcular a porcentagem de explicação de cada atributo na estimação dos modelos. Observa-se que, em quase todas as diretorias, o atributo 1 explicou melhor o modelo, exceto na diretoria 8, onde tanto o atributo 1 como o atributo 2 têm, aproximadamente, a mesma porcentagem de explicação para o modelo.

Assim, nas diretorias 1, 2, 4, 5, 7, 8, 10, 11 e 12 observa-se que a maioria dos indivíduos deram maior importância aos estímulos 3, 7, 6 e 5. Isto é, Solidez da instituição e Parceria com o relacionamento com empresas são os estímulos que a entidade financeira deve priorizar, seguidos pela Agilidade operacional da estrutura administrativa e Autonomia dos profissionais. Percebe-se que estes dois últimos estímulos podem ser considerados como Atendimento ou Serviço da entidade financeira para seus clientes. Neste gradiente, os estímulos Especialização em grandes clientes, Liderança de mercado e Localização Geográfica das agências são os menos importantes.

Por outro lado, nas diretorias 3, 6 e 9 percebe-se a formação de grupos, isto é, os indivíduos foram agrupados devido a algum fator comum entre eles. Estes gráficos podem ajudar a segmentar o mercado. Na diretoria 3, percebe-se que 22 indivíduos localizam suas preferências mais acentuadamente com a parceria no relacionamento com empresas, enquanto que 19 indivíduos comparecem localizando maior preferência com a agilidade operacional da estrutura administrativa e autonomia dos profissionais. Na diretoria 6, observa-se que 22 indivíduos localizam suas preferências mais concentrada em relação a autonomia dos profissionais e especialização em grandes clientes e 18 indivíduos comparecem localizando suas preferências mais acentuadamente na parceria no relacionamento com empresas e agilidade operacional da estrutura administrativa. Na diretoria 9, observa-se também que 17 indivíduos estão mais preocupados com a parceria no relacionamento com empresas, enquanto que 18 indivíduos estão preocupados com agilidade operacional da estrutura administrativa e autonomia dos profissionais.

5.10. Resultados Gerais da Aplicação.

De acordo com a metodologia empregada neste trabalho e com base nos resultados obtidos, pode-se chegar às seguintes conclusões gerais da aplicação:

1.- A metodologia mostrou-se adequada, fornecendo resultados coerentes, tanto na aplicação do MDS assim como na aplicação dos Modelos de Preferência. Os resultados obtidos da combinação dessas duas técnicas multivariadas foram aceitos pelas empresas envolvidas. Além disso, proporcionaram informação para o melhor atendimento a seus clientes.

2.- A solução do MDS representou graficamente a ordem dos estímulos, isto é, observamos através dos gráficos que o comportamento dos estímulos pode ser entendido pelo seu grau de preferência, regularmente manifestado através dos grupos.

3.- Os Modelos de Preferência via MDS serviram para observar o posicionamento dos indivíduos frente aos estímulos. Isto ajudou a visualizar graficamente o comportamento de cada indivíduo no espaço dos estímulos, bem como agrupamentos dos indivíduos que quando rastreados e entendidos, tornam-se ferramentas úteis na segmentação de mercado.

4.- Os Modelos de Preferência mostraram a porcentagem de explicação de cada atributo (ou dimensão) na estimação dos modelos, o que reflete distintas percepções dos indivíduos dentro dos grupos quanto á importância relativa entre os atributos.

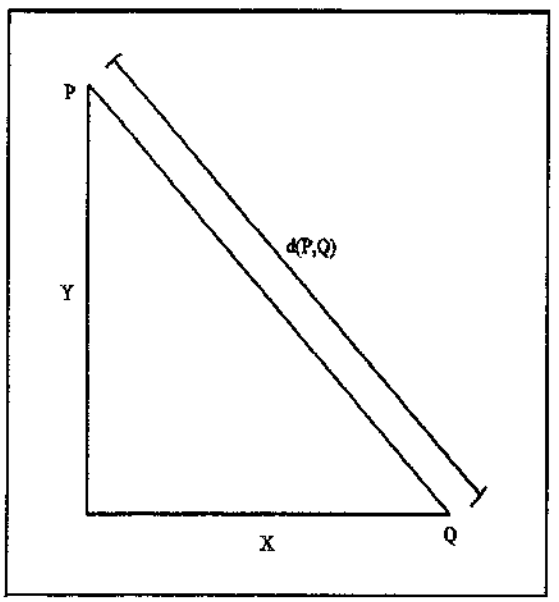
5.- Apesar de haver softwares específicos para o MDS e para os Modelos de Preferência, usamos o software disponível S.A.S., que se mostrou eficiente para a maioria das etapas, no entanto na parte gráfica recorremos a outros softwares. Foram usados além do S.A.S. o software estatístico MINITAB (parte gráfica), e os processadores de gráficos e textos WORDPERFECT e MICROSOFT WORD.

Apêndice A

DISTÂNCIAS E SIMILARIDADES

A.1. Definição.

Dados dois pontos $P \in R^p$ e $Q \in R^p$ com representação vetorial X e Y.



A função d(P,Q), definida com a distância entre P e Q possui as seguintes propriedades:

i) Simetria: d(P,Q)=d(Q,P).
ii) d(P,Q)>0 é não negativa (quando P≠Q).
iii) Sinal de Identificação: d(P,P)=0.

iv) $d(P,Q)=0 \Leftrightarrow P=Q$.

v) Satisfaz a desigualdade triangular: $d(P,Q) \leq d(P,R)+d(R,Q)$.

OBSERVAÇÕES

- Se d satisfaz às cinco propriedades acima é chamado MÉTRICO.
- Se d não satisfaz à desigualdade triangular, é chamado SEMI-MÉTRICO.

A.2. Análise Para Dados Quantitativos.

Dados os pontos $X^t=(x_1,x_2,.....x_n)\in R^n$ e $Y^t=(y_1,y_2,.....,y_n) \in R^n$, a distância entre X e Y definida por:

$$d_{(X,Y)} = \left[\sum_{i=1}^{n}\left(x_i - y_i\right)^p\right]^{1/p} \qquad (a1)$$

é chamada de distância de MINKOWSKI ou simplesmente Lp, que satisfaz às 5 propriedades mencionadas acima, por tanto $d_{(X,Y)}$ é métrico. Temos os seguintes casos especiais de interesse:

Se p=1 na equação (a1) reduz-se a:

$$d_{(X,Y)} = \left[\sum_{i=1}^{n}\left|x_i - y_i\right|\right] \qquad (a2)$$

chamada distância CITY_BLOCK, que satisfaz as 5 propriedades de distância, portanto é métrica.

Se fixamos p=2, a equação (a1) reduz-se a:

$$d_{(X,Y)} = \left[\sum_{i=1}^{n}\left(x_i - y_i\right)^2\right]^{1/2} = \|X-Y\|^2 \qquad (a3)$$

conhecida como distância EUCLIDIANA, que satisfaz as 5 propriedades de distância portanto é métrica.

Quando $p \rightarrow \infty$ a distância de Minkowski converge para:

$$d_{(X,Y)} = \max_{1\leq j\leq n}\left|x_j - y_j\right| \qquad (a4)$$

A.3. Distância Euclidiana de Amostras Aleatórias.

A distância Euclidiana entre vetores do R^n , quando estes representam amostras aleatórias de tamanho n tem expresso em função das variância e covariância correspondentes. Para a distância entre vetores padronizados a expressão envolve o coeficiente de correlação como mostramos a seguir.

A.3.1. Caso Particular (2 variáveis)

Suponha duas amostras, $X^t=(x_1,x_2,.....x_n)\in R^n$ e $Y^t=(y_1,y_2,.....,y_n)\in R^n$, onde X e Y são variáveis aleatórias, considere suas médias amostrais:

$$\bar{x}=\frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{y}=\frac{\sum_{i=1}^{n} y_i}{n} \tag{a5}$$

Considere ainda as variâncias amostrais de X e Y, como também a covariância entre eles,

$$S_X^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \qquad S_Y^2=\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1} \tag{a6}$$

$$S_{XY}^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n-1} \tag{a7}$$

Definamos X^* e Y^* como os respectivos **VETORES DE RESÍDUOS**, ou seja:

$$X^*=\begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ \vdots \\ x_n^* \end{pmatrix}=\begin{pmatrix} x_1-\bar{x} \\ x_2-\bar{x} \\ \vdots \\ \vdots \\ x_n-\bar{x} \end{pmatrix}_{nx1} \qquad Y^*=\begin{pmatrix} y_1^* \\ y_2^* \\ \vdots \\ \vdots \\ y_n^* \end{pmatrix}=\begin{pmatrix} y_1-\bar{y} \\ y_2-\bar{y} \\ \vdots \\ \vdots \\ y_n-\bar{y} \end{pmatrix}_{nx1} \tag{a8}$$

Observa-se que os vetores X^* e Y^* são transladados à origem. Os vetores X^* e Y^* são denominados de **desvio escore**, Takeuchi, Yanai e Mukherjee (1984). A distância euclidiana entre os vetores de resíduos X^* e Y^* é:

$$
\begin{aligned}
d^2_{(X^*,Y^*)} &= \langle X^* - Y^*, X^* - Y^* \rangle \\
&= \left\| X^* - Y^* \right\|^2 \\
&= \sum_{i=1}^{n} \left(x_i^* - y_i^* \right)^2 \\
&= \left(x_1^* - y_1^* \right)^2 + \left(x_2^* - y_2^* \right)^2 + \ldots\ldots\ldots\ldots + \left(x_n^* - y_n^* \right)^2 \\
&= \sum_{i=1}^{n} \left(x_i^{*2} - 2x_i^* y_i^* + y_i^{*2} \right) \\
&= \sum_{i=1}^{n} x_i^{*2} - 2\sum_{i=1}^{n} x_i^* y_i^* + \sum_{i=1}^{n} y_i^{*2} \\
&= \langle X^*, X^* \rangle - 2\langle X^*, Y^* \rangle + \langle Y^*, Y^* \rangle \\
&= \sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2 - 2\sum_{i=1}^{n} \left(x_i - \bar{x} \right)\left(y_i - \bar{y} \right) + \sum_{i=1}^{n} \left(y_i - \bar{y} \right)^2 \\
&= (n-1)S_X^2 - 2(n-1)S_{XY}^2 + (n-1)S_Y^2 \\
&= (n-1)\left[S_X^2 - 2S_{XY}^2 + S_Y^2 \right]
\end{aligned}
$$

Por conseguinte temos que:

$$
d^2_{(X^*,Y^*)} = (n-1)\left[S_X^2 - 2S_{XY}^2 + S_Y^2 \right] \qquad \text{(a9)}
$$

Observa-se que a distância Euclidiana entre os vetores de resíduos X^* e Y^* pode ser expressa em função das variâncias estimadas de X (S_X^2), e de Y (S_Y^2) e a covariância estimada entre X e Y(S_{XY}^2). Observa-se que a distância Euclidiana mede de certo modo o grau de relação entre os vetores X^* e Y^*.

Vejamos o que acontece se padronizamos os vetores X e Y da seguinte maneira:

$$Z_X = \begin{pmatrix} z_{x_1} \\ z_{x_2} \\ \vdots \\ z_{x_n} \end{pmatrix} = \begin{pmatrix} \frac{x_1 - \bar{x}}{S_X^2} \\ \frac{x_2 - \bar{x}}{S_X^2} \\ \vdots \\ \frac{x_n - \bar{x}}{S_X^2} \end{pmatrix}_{nx1} \qquad Z_Y = \begin{pmatrix} z_{y_1} \\ z_{y_2} \\ \vdots \\ z_{y_n} \end{pmatrix} = \begin{pmatrix} \frac{y_1 - \bar{y}}{S_Y^2} \\ \frac{y_2 - \bar{y}}{S_Y^2} \\ \vdots \\ \frac{y_n - \bar{y}}{S_Y^2} \end{pmatrix}_{nx1} \tag{a10}$$

Calculando a distância Euclidiana de Z_X e Z_Y temos:

$$\begin{aligned} d^2_{(z_x, z_y)} &= \sum_{i=1}^{n} \left(z_{x_i} - z_{y_i} \right)^2 \\ &= \sum_{i=1}^{n} z^2_{x_i} - 2 \sum_{i=1}^{n} z_{x_i} z_{y_i} + \sum_{i=1}^{n} z^2_{y_i} \\ &= \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{S_x} \right)^2 - 2 \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{S_x} \right)\left(\frac{y_i - \bar{y}}{S_y} \right) + \sum_{i=1}^{n} \left(\frac{y_i - \bar{y}}{S_y} \right)^2 \\ &= (n-1) - 2\frac{(n-1)\mathrm{cov}(x,y)}{S_x S_y} + (n-1) \\ &= 2(n-1)\left[1 - \frac{\mathrm{cov}(x,y)}{\sqrt{S_x^2}\sqrt{S_y^2}} \right] \\ &= 2(n-1)\left[1 - R_{XY} \right] \end{aligned}$$

Consequentemente temos que:

$$d^2_{(Z_X, Z_Y)} = 2(n-1)\left[1 - R_{XY} \right] \tag{a11}$$

onde: $\mathbf{R_{XY}}$ é o coeficiente de correlação amostral de Pearson entre X e Y.

Observa-se que a distância Euclidiana está relacionada com o coeficiente de correlação de Pearson, então concluimos que a distância Euclidiana além de medir magnitude também mede o grau de relação linear entre indivíduos, observações ou objetos.

COMENTÁRIOS ADICIONAIS

Sabe-se que o coeficiente da correlação de Pearson ($R_{X,Y}$) varia de -1 a 1, isto é: $1 \leq R_{XY} \leq -1$. Analisando os extremos da correlação, temos o seguinte: se $R_{X,Y}$ é igual ou próximo a 1, a distância Euclidiana entre X e Y é zero ou próximo de zero, indicando que quanto mais X e Y estão relacionados linearmente a distância entre eles se aproxima de zero. Analogamente, correlação próxima de zero indicam maior distância entre X e Y.

A.3.2. Caso Geral (p variaveis).

Seja uma matriz de dados X_{nxp} formado por n observações ou indivíduos e medida em p variáveis; definida da seguinte maneira:

$$X_{nxp} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{matrix} \vec{X}_1^t \\ \vec{X}_2^t \\ \vdots \\ \vec{X}_n^t \end{matrix} \qquad \text{(a12)}$$

A partir da matriz de dados considere o vetor de médias, a matriz de covariâncias e a matriz de correlações isto é:

Vetor de médias :

$$\vec{\bar{x}} = \frac{\sum_{i=1}^{n} \vec{X}_i}{n} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix}_{Px1} \qquad \text{(a13)}$$

Matriz de variância e covariância :

$$S = \frac{\sum_{i=1}^{n} \left(\vec{X}_i - \vec{\bar{x}}\right)\left(\vec{X}_i - \vec{\bar{x}}\right)^t}{n-1} \tag{a14}$$

Calculando a distância Euclidiana entre os vetores $\vec{X}_i$ e $\vec{X}_j$ temos:

$$d_{ij}^2 = \sum_{k=1}^{p} \left(x_{ik} - x_{jk}\right)^2 = \left\|\vec{X}_i - \vec{X}_j\right\|^2 \tag{a15}$$

Para efeitos de cálculo vamos utilizar o seguinte fato: dado o vetor $\mathbf{e_i}$, de dimensão n, com o i-ésimo elemento igual a 1 e 0 nas demais posições isto é:

$$e_i^t = (0,0,...,1,...,0)_{1xn} \tag{a16}$$

Denotamos o i-ésimo vetor linha de X_{nxp} por $\vec{X}_i$ por:

$$\vec{X}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix}_{px1} = X^t\left(e_i\right) \tag{a17}$$

Substituindo a equação (a17) em (a15) temos:

$$\begin{aligned} d_{ij}^2 = d_X^2\left(e_i, e_j\right) &= \left\|\vec{X}_i - \vec{X}_j\right\|^2 \\ &= \left(\vec{X}_i - \vec{X}_j\right)^t \left(\vec{X}_i - \vec{X}_j\right) \\ &= \left[X^t\left(e_i\right) - X^t\left(e_j\right)\right]^t \left[X^t\left(e_i\right) - X^t\left(e_j\right)\right] \\ &= \left(e_i - e_j\right)^t XX^t\left(e_i - e_j\right) \end{aligned}$$

isto é:

$$d_{ij}^2 = d_{\vec{X}}^2\left(e_i, e_j\right) = \left(e_i - e_j\right) XX^t \left(e_i - e_j\right)^t \tag{a18}$$

Definamos a **Matriz de Resíduos** da seguinte maneira:

$$X^*_{nxp} = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{12} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \cdots & x_{np} - \bar{x}_p \end{pmatrix} \begin{matrix} \vec{X}^{*t}_1 \\ \vec{X}^{*t}_2 \\ \vdots \\ \vec{X}^{*t}_n \end{matrix} \tag{a19}$$

O i-ésimo indivíduo pode ser escrito da seguinte forma:

$$\vec{X}^*_i = \begin{pmatrix} x_{i1} - \bar{x}_1 \\ x_{i2} - \bar{x}_2 \\ \vdots \\ \vdots \\ x_{ip} - \bar{x}_p \end{pmatrix}_{px1} = X^{*t}\left(e_i\right) \tag{a20}$$

Outra forma de escrever a matriz (a19) é:

$$X^*_{nxp} = P^{\perp}_M X_{nxp} \tag{a21}$$

onde:

$$P^{\perp}_M = \left(I_{nxn} - \mathbf{1}_{nx1}\mathbf{1}^t_{nx1}\right) \tag{a22}$$

$$I_{nxn} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}_{nxn} \quad \mathbf{1}^t = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}_{1xn} \tag{a23}$$

Então a distância Euclidiana dos pontos $\vec{X}^*_i$ e $\vec{X}^*_j$ de X^* ,usando (a20) é:

$$
\begin{aligned}
d^2_{X^*}(e_i,e_j) &= \left\|\vec{X}^*_i - \vec{X}^*_j\right\|^2 \\
&= \left(\vec{X}^*_i - \vec{X}^*_j\right)^t\left(\vec{X}^*_i - \vec{X}^*_j\right) \\
&= \left[X^{*t}(e_i) - X^{*t}(e_j)\right]^t\left[X^{*t}(e_i) - X^{*t}(e_j)\right] \\
&= (e_i - e_j)^t X^* X^{*t}(e_i - e_j) \\
&= (e_i - e_j)^t P^{\perp}_M X X^t P^{\perp}_M (e_i - e_j) \\
&= (e_i - e_j)^t X X^t (e_i - e_j) \\
&= d^2_X(e_i,e_j)
\end{aligned}
$$

então:

$$d^2_X(e_i,e_j) = d^2_{X^*}(e_i,e_j) \tag{a24}$$

Está provado que a distância Euclidiana entre dois vetores (i,j) da matriz X é a mesma que a distância entre os vetores (i,j) da matriz X^*, isto é **A distância Euclidiana é invariante a mudança de origem.**

Por outro lado, seja a seguinte transformação: $X_W=XW$, onde W é uma matriz ortogonal de ordem r, temos que a distância Euclidiana entre 2 linhas da matriz X_W pode ser expressa da seguinte forma:

$$
\begin{aligned}
d^2_{X_W}(e_i,e_j) &= (e_i,e_j)^t XWW^t X^t (e_i,e_j) \\
&= (e_i,e_j)^t XX^t (e_i,e_j) \\
&= d^2_X(e_i,e_j)
\end{aligned}
$$

então:

$$d^2_{X_W}(e_i,e_j) = d^2_X(e_i,e_j) \tag{a25}$$

Assim provamos que a **distância Euclidiana é invariante a qualquer transformação ortogonal dos eixos.**

A.4. Outros Tipos de Distância.

Nesta seção apresentamos outros tipos de distância.

Distância De Karl Pearson

A distância ao quadrado de Karl Pearson entre os pontos $\vec{X}_i$ e $\vec{X}_j$ é definida como:

$$d_{ij}^2 = \sum_{k=1}^{p} \frac{\left(x_{ik} - x_{jk}\right)^2}{S_k^2} \qquad \text{(a26)}$$

onde S_k^2 é a variância da k-esima variável.

Distância De Mahalanobis

A distância quadrada de Mahalanobis, entre os pontos $\vec{X}_i$ e $\vec{X}_j$, é definido por:

$$d_{ij}^2 = \left(\vec{x}_i - \vec{x}_j\right)^t \hat{\Sigma}^{-1} \left(\vec{x}_i - \vec{x}_j\right) \qquad \text{(a27)}$$

onde Σ é matriz de covariâncias. Havendo estrutura de grupos, a distância de Mahalanobis pode ser definida em função do vetor de médias de cada grupo, isto é:

$$d_{ij}^2 = \left(\vec{\bar{x}}_i - \vec{\bar{x}}_j\right)^t \hat{\Sigma}^{-1} \left(\vec{\bar{x}}_i - \vec{\bar{x}}_j\right) \qquad \text{(a28)}$$

Distância Qui-quadrada

A distância Qui-quadrado pode ser interpretada como uma distância Euclidiana ponderada. Esta distância geralmente é aplicada sobre tabelas de contingência. A continuação mostramos a forma de calcular, para isto utilizaremos a seguinte tabela de contingência.

$$F = \begin{array}{c} \\ R_1 \\ R_2 \\ \vdots \\ R_q \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & \cdots & E_p \end{array} \\ \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1p} \\ f_{21} & f_{22} & \cdots & f_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ f_{q1} & f_{q2} & \cdots & f_{qp} \end{bmatrix} \\ \begin{array}{cccc} f_{.1} & f_{.2} & \cdots & f_{.p} \end{array} \end{array} \begin{array}{c} \\ f_{1.} \\ f_{2.} \\ \vdots \\ f_{q.} \\ \sum_{i=1}^{q}\sum_{j=1}^{p} f_{ij} \end{array} \qquad (a29)$$

onde:

f_{ij} : Indica as frequências observadas.

$f_{i.} = \sum_{r=1}^{p} f_{ri}$: Indica a marginal linha .

$f_{.j} = \sum_{s=1}^{q} f_{js}$: Indica a marginal coluna.

Para todo i=1,2,...,q. e j=1,2,...,p.

A distância Qui-quadrada entre as colunas E_u e E_v da matriz (a29) pode ser calculada da seguinte forma:

$$d^2_{E_u E_v} = (\vec{e}_u - \vec{e}_v)^t D^{-1} (\vec{e}_u - \vec{e}_v) \qquad (a30)$$

onde:

$$\vec{e}_u = \left(\frac{f_{1u}}{f_{.u}}, \frac{f_{2u}}{f_{.u}}, \ldots, \frac{f_{qu}}{f_{.u}}\right)^t \quad \vec{e}_v = \left(\frac{f_{1v}}{f_{.v}}, \frac{f_{2v}}{f_{.v}}, \ldots, \frac{f_{qv}}{f_{.v}}\right)^t \quad \text{para todo u,v=1,2,...,p.}$$

D é uma matriz diagonal de ponderação definida da seguinte maneira:

$$D = \begin{pmatrix} \frac{f_{1.}}{\sum_{i=1}^{q}\sum_{j=1}^{p} f_{ij}} & 0 & \cdots & 0 \\ 0 & \frac{f_{2.}}{\sum_{i=1}^{q}\sum_{j=1}^{p} f_{ij}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{f_{q.}}{\sum_{i=1}^{q}\sum_{j=1}^{p} f_{ij}} \end{pmatrix} \qquad (a31)$$

Analogamente, pode ser calculada a distância Qui-quadrada entre dois categorias de R.

A.5. Coeficientes de Similaridade

Os coeficientes de similaridade, às vezes, conhecidos como "medida de distância" ou "dissimilaridade", referem-se basicamente a medida de similaridade entre 2 pontos A e B.

A.5.1. Definição.

A medida de similaridade entre 2 pontos A e B, definida por S(A,B), satisfaz as seguintes propriedades:

i) S(A,B)=S(B,A) (simetria)

ii) S(A,B)>0 (sempre é positiva)

iii) S(A,B) aumenta, se a similaridade entre os pontos A e B aumenta (monotonia).

Os coeficientes de similaridade são usados quando queremos medir o grau de semelhança ou diferença entre 2 objetos.

A.5.2. Análise para variáveis Qualitativas

Suponhamos que as variáveis $x_1,x_2,...,x_p$ são dicotômicas, baseadas em ausência(-) ou presença (+) de caracteres qualitativos. Um indivíduo I_i fica caracterizado pelas presenças ou ausências de n caracteres, por exemplo:

	x_1	x_2	x_3	...	x_4
I_i	+	-	-	...	+

A informação útil é, então, o número de caracteres presentes sobre os n caracteres estudados. Para estudar a associação entre dois indivíduos I_1 e I_2, constrói-se a seguinte tabela de freqüências:

		I_i	
		+	-
I_j	+	a	b
	-	c	d

onde :

n = a+b+c+d

a : Indica o número de caracteres comuns presentes,

d : Indica o número de caracteres comuns ausentes, etc.

Mede-se a associação através de um coeficiente de similaridade s_{ij}, definido como função de a,b, e c :

$$s_{ij} = f(a,b,c)$$

tal que:

1) s_{ij} é crescente em a;

2) s_{ij} é decrescente em b e c;

3) s_{ij} é simétrica em b e c.

O coeficiente de similaridade s_{ij} como seu nome indica, é uma medida do grau de semelhança ou similaridade entre I_i e I_j em relação a os n caracteres. A maioria dos coeficientes de similaridade varia entre 0 e 1, sendo:

$s_{ij} = 0$, se c+b=n

$s_{ij} = 1$, se a+d=n

s_{ij}=0 quando todo caracter presente em I_i não está presente em I_j (discrepância ou dissimilaridade total) e s_{ij}=1 quando tudo caracter presente I_i está presente também em I_j (similaridade total).

Vários autores propuseram coeficientes de similaridades com essas propriedades:

- **Sokal e Michener** : $\frac{a+d}{n}$

- **Sokal e Sneath** : $\frac{a}{a+2(b+c)}$

- **Jaccard** : $\frac{a}{a+b+c}$

- **Rusell e Rao** : $\frac{a}{n}$

- Dice : $\frac{2d}{(a+b)(a+c)}$

- Kulezynski : $\frac{a}{b+c}$

- Harman : $1-\frac{2(b+c)}{n}$

- Roger e Tanimoto : $\frac{n-(b+c)}{n+(b+c)}$

- Ochiai : $\frac{a}{\sqrt{a+b}\sqrt{a+c}}$

Existem outros coeficientes que expressam dependência estocástica entre I_i e I_j e variam de -1 a +1. A dissimilaridade total corresponde a -1 e a similaridade total a +1. O valor zero identifica-se com a noção de independência estocástica. Exemplos de estes coeficientes são:

- Yule : $\frac{ad-bc}{ad+bc}$

- Pearson : $\frac{ad-bc}{[(a+c)(b+d)(a+b)(c+d)]^{1/2}}$

Uma vez definido um coeficiente de similaridade, este é calculado para todos os pares de indivíduos I_i e I_j , e o resultado disposto em uma matriz simétrica, da seguinte maneira:

$$S=\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix}$$

é chamada de matriz de similaridades entre os indivíduos.

Não existe um critério absoluto que permita decidir o coeficiente de similaridade mais adequado. Na eleição de um determinado coeficiente intervém, o tipo de dados que serão

representados e o peso que se dará as freqüências a, b, c, d. Trata-se de um problema que deve ser resolvido para cada situação experimental concreta. No entanto, é conveniente ter em conta os seguintes critérios (mais importantes):

a) s_{ij} não deve ser função de d, pois aumentando caracteres arbitrários não comuns, poderiam fazer falsamente similares indivíduos que não são.

b) O coeficiente deve ser tal que a matriz S seja semidefinida ou definida positiva.

c) O coeficiente deve satisfazer (aproximadamente) a propriedade métrica.

Apêndice B

DEFINIÇÃO DO MÉTODO "Q" E MÉTODO "R" NA ANÁLISE MULTIVARIADA

B.1. Definição

A definição dos métodos Q e R na análise multivariada, deve-se basicamente à forma como analisamos uma matriz de dados. Na maioria das vezes estamos interessados em trabalhar sobre as variáveis, e outras vezes podemos estar interessados em trabalhar sobre os indivíduos, a maioria das técnicas multivariadas, trabalham sobre as variáveis.

A diferença básica entre estes dois termos é a seguinte: A partir de uma matriz de dados X_{nxp} podemos calcular a matriz de covariâncias entre as variáveis como também uma matriz de relações e interdependências entre os indivíduos. As técnicas Multivariadas que trabalham utilizando a matriz de covariâncias (ou correlação) denomina-se Método R. As técnicas multivariadas que trabalham usando basicamente matriz de relações e interdependências dos indivíduos denomina-se Método Q. O presente apêndice será desenvolvido baseado no trabalho de Jollife (1986) e Jackson (1991).

B.2. Semelhanças e diferenças entre estes dois métodos.

Seja uma matriz de dados X_{nxp} formada por n observações (ou indivíduos) e medida em p variáveis ($p<n$); como definida em (n1). Esta matriz pode ser transladada à origem da seguinte maneira:

$$Z_{nxp} = X_{nxp} - 1_{nx1}\bar{\bar{x}}^{t}_{1xp} \qquad (b1)$$

Sabe-se que a matriz de covariâncias das variáveis é :

$$S^2 = \frac{Z^tZ}{n-1} \qquad (b2)$$

Da mesma forma, a matriz de relações e interdependências dos indivíduos é definida por:

$$S@2 = \frac{ZZ^t}{n-1} \qquad (b3)$$

Então observa-se a diferença básica está nas matrizes $(Z^tZ)_{pxp}$ e $(ZZ^t)_{nxn}$.

As seguintes propriedades podem ser mostradas:

- A matriz $(Z^tZ)_{pxp}$ é uma matriz Definida Positiva.
- A matriz $(ZZ^t)_{nxn}$ é uma matriz Semi-definida Positiva.
- Os traços de ambas matrizes são iguais isto é:

$$Tr\left[\left(Z^tZ\right)_{pxp}\right] = Tr\left[\left(ZZ^t\right)_{nxn}\right] \qquad (b4)$$

- Os postos das matrizes são iguais a "p", isto é:

$$Posto\left[\left(Z^tZ\right)_{pxp}\right] = Posto\left[\left(ZZ^t\right)_{nxn}\right] = p \quad \text{se: } n>p \qquad (b5)$$

Pelo teorema de "Descomposição Espectral" ou teorema de "Descomposição de Jordan"; Mardia, Kent e Bibby (1979), temos que:

A matriz simétrica $(Z^tZ)_{pxp}$ pode ser escrita como:

$$\left(Z^tZ\right)_{pxp} = \Gamma\Lambda\Gamma^t = \sum_{i=1}^{p} \lambda_i \gamma_{(i)}\gamma_{(i)}^t \qquad (b6)$$

onde: Λ é uma matriz diagonal formada pelos autovalores $(\lambda_1,\lambda_2,...,\lambda_p)$ da matriz $(Z^tZ)_{pxp}$ e Γ uma matriz ortogonal cujas colunas são autovetores ortonormais correspondentes aos autovalores.

Do mesmo modo pode-se escrever a matriz $(ZZ^t)_{nxn}$ como:

$$\left(ZZ^t\right)nxn = \Gamma^{@}\Lambda^{@}\Gamma^{@t} = \sum_{i=1}^{n} \lambda_i^{@}\gamma_{(i)}^{@}\gamma_{(i)}^{@t} \qquad (b7)$$

onde: $\Lambda^{@}$ é uma matriz diagonal formada pelos autovalores $(\lambda_1^{@},\lambda_2^{@},.....,\lambda_p^{@},\lambda_{p+1}^{@},.....,\lambda_n^{@})$ da matriz $(ZZ^t)_{nxn}$ e $\Gamma^{@}$ uma matriz ortogonal, cujas colunas são autovetores padronizados correspondentes aos autovalores.

Pode-se mostrar que :

$$\begin{aligned}
\lambda_1 &= \lambda_1^{@} \\
\lambda_2 &= \lambda_2^{@} \\
&\vdots \\
\lambda_p &= \lambda_p^{@} \\
\lambda_{p+1}^{@} &= 0 \\
&\vdots \\
\lambda_n^{@} &= 0
\end{aligned} \qquad (b8)$$

Então, os autovetores $\gamma_{(i)}$ e $\gamma_{(i)}^{@}$ são proporcionais, pois estão na mesma direção (ambas têm o mesmo autovetor).

Deste modo a técnica de Scaling Multidimensional (MDS) pode ser considerada como um método "Q" ou "R", pois, MDS pode trabalhar sobre os indivíduos ou sob as variáveis.

Apêndice C

ESTRUTURA DE MATRIZES

C.1. Introdução

Neste apêndice, estamos interessados na aproximação de matrizes; formalizaremos as características destas aproximações. Uma das principais características é que as matrizes aproximadas têm posto menor que as matrizes originais. Eckart e Young (1936) e Young e Householder (1938) apresentam descrições pioneiras de aproximações de matrizes no contexto de Autoestrutura de matrizes. Por outro lado artigos sobre aproximação de matrizes com caraterísticas particulares são descritos por Gabriel e Zamir (1976). Também Greenacre e Underhill (1982) descrevem de uma maneira simplificada aproximações de matrizes, que são a base para desenvolver este apêndice.

Serão descritos, brevemente, conceitos matemáticos de álgebra linear, como descomposição de valores singulares (DVS), descomposição de valores singulares generalizados (DVSG), aproximações de matrizes com menor posto, qualidade e erro de aproximação de matrizes, para então abordamos o tema sobre autoestrutura de uma matriz produto escalar e finalmente relacionar produto escalar com distância, baseados na ideia de Young e Householder (1938).

C.2. Descomposição de Valores Singulares (DVS) ou Estrutura Básica.

Seja A uma matriz semi-definida positiva de ordem (nxp) retangular e de posto (r). Pela teoria da álgebra matricial, sabe-se que a matriz A pode ser descomposto em 3 elementos de Estrutura Simples:

$$A_{nxp} = U_{nxr} D\lambda_{rxr} V^t_{rxp} \qquad (c1)$$

onde:

$D\lambda$: é uma matriz diagonal com elementos positivos $\lambda_1, \lambda_2, \ldots\ldots, \lambda_r$.

U e V são matrizes com colunas ortonormais, isto é, $U^t U=I$ e $V^t V=I$.

Se A é uma matriz simétrica e quadrada então U=V. Greenacre e Leslie denominam (c1) de Estrutura Básica, e se A é simétrica, (c1) é denominada simplesmente Autoestrutura da matriz A.

C.3. Aproximações de Matrizes com Menor Posto.

O valor prático da Descomposição de Valores Singulares é que permite que uma matriz A, geralmente de posto maior (r) seja aproximada por uma matriz $\hat{A}_{[k]}$ de posto menor (k), isto é, (k<r). Por outro lado como os autovalores de A são ordenados isto é: $\lambda_1 > \lambda_2 > \ldots\ldots > \lambda_r$ (pois A é semi-definida positiva), então se $U_{[k]}$ e $V_{[k]}$ contem os primeiros k autovetores de U e V (ordenada correspondentemente) e $D\lambda_{[k]}$ uma matriz diagonal formado pelos primeiros k autovalores, então uma aproximação matricial é definida como:

$$\hat{A}_{[k]nxp} = U_{[k]nxk} D\lambda_{[k]kxk} V^t_{[k]kxp} \qquad (c2)$$

onde $\hat{A}_{[k]}$ é uma aproximação de A, no sentido que dentre todas as matrizes $A_{[k]}$ de posto k, ela é que minimiza a equação abaixo:

$$\left\| A - A_{[k]} \right\|^2 = tr\left[\left(A - A_{[k]} \right) \left(A - A_{[k]} \right)^t \right] \qquad (c3)$$

Mardia, Kent e Bibby (1979) baseia-se na equação (c3) para achar uma medida de qualidade de ajuste para a parte de Escalonamento métrico.

C.4. Descomposição de Valores Singulares generalizados.(DVSG)

A Descomposição de valores singulares generalizados (DVSG) é uma generalização da Descomposição de valores singulares (DVS), no qual a técnica de MDS é baseada. Greenacre e Leslie a denominam de Autoestrutura generalizada.

A norma de uma matriz é generalizada da seguinte forma :

$$\|A\|^2_{\Omega,\Phi} = tr\left[\Omega A \Phi A^t\right] \tag{c4}$$

onde : Ω,Φ são matrizes simétricas definidas positivas.

Pode-se notar que na equação (c3) a norma é um caso especial quando $\Omega=\Phi=\mathbf{I}$.

Como a expressão (c4) é função da traço, também pode ser escrita assim:

$$\|A\|^2_{\Omega,\Phi} = \sum_{i=1}^{n}\sum_{i^\circ=1}^{n} \omega_{ii^\circ} a_i^t \Phi a_{i^\circ} \tag{c5}$$

ou ainda:

$$\|A\|^2_{\Omega,\Phi} = \sum_{j=1}^{p}\sum_{j^\circ=1}^{p} \phi_{jj^\circ} a_j^t \Omega a_{j^\circ} \tag{c6}$$

Na equação (c5) observa-se que a norma quadrada é uma soma ponderada por ω_{ii} de todos os produtos escalares entre as linhas a_i de A na métrica Φ. Analogamente na expressão (c6) observa-se que a norma quadrada é uma soma ponderada por ϕ_{jj} de todos os produtos escalares entre as colunas a_j de A na métrica de Ω.

Um caso especial, que vamos usar mais adiante, ocorre quando qualquer dos dois Ω ou Φ é uma matriz diagonal. Por exemplo, se $\Omega=D\omega$, então a expressão (c5) é reduzida a:

$$\|A\|^2_{D\omega,\Phi} = \sum_{i=1}^{n} \omega_i \|a_i\|^2 \Phi \tag{c7}$$

que é uma soma ponderada por ω_i das normas quadradas dos vetores linhas a_i de A, na métrica de Φ.

Por outro lado a matriz A pode ser descomposta em uma forma mais geral do que DVS, da seguinte maneira:

$$A = ND_{\alpha}M^{t} \tag{c8}$$

onde:

$\mathbf{N^{t}\Omega N=I}$ e $\mathbf{M^{t}\Phi M=I}$.

A expressão (c8) é conhecido como Descomposição de Valores Singulares Generalizados (DVSG). Nota-se que os vetores básicos N e M são ortogonais nas métricasde Ω e Φ, respectivamente.

Então como no caso da equação (c2) uma aproximação matricial de A pode ser feita da seguinte forma:

$$\hat{A}_{[k]} = N_{[k]}D\alpha_{[k]}M_{k}^{t} \tag{c9}$$

onde $\hat{A}_{[k]}$ é uma aproximação de A, no sentido em que dentre todas as matrizes $A_{[k]}$ de posto k, ela é a que minimiza a expressão:

$$\left\|A - A_{[k]}\right\|_{\Omega,\Phi}^{2} = \mathrm{tr}\left[\Omega\left(A - A_{[k]}\right)\Phi\left(A - A_{[k]}\right)^{t}\right]$$

C.5. Consequências da Descomposição de Valores Singulares Generalizados.

i) Na prática a estrutura básica (c1) é obtida a través da autoestrutura, é dizer da descomposição de autovalores e autovetores.

ii) Para o problema geral na equação (c8) temos:

$$A = ND_{\alpha}M^{t}$$

onde:

$N^{t}\Omega N=I$ e $M^{t}\Phi M=I$.

desta forma, temos :

$$
\begin{aligned}
A\Phi A^t &= ND_\alpha M^t \Phi A^t \\
&= ND_\alpha M^t \Phi \left[ND_\alpha M^t \right]^t \\
&= ND_\alpha M^t \Phi M D_\alpha^t N^t \\
&= ND_\alpha I D_\alpha^t N^t \\
&= ND_\alpha^2 N^t
\end{aligned}
$$

então:

$$A\Phi A^t = ND_\alpha^2 N^t \qquad \text{(c10)}$$

onde $N^t\Omega N=I$ desde que: $M^t\Phi M=I$

A expressão (c10) é conhecido como: "Autoestrutura generalizada". Os autovetores N são os vetores básicos esquerdos de A (é dizer autovetores esquerdos de A) e os autovalores são os valores básicos quadrados de A.

Note que $A\Phi A^t$ é uma matriz produto escalar entre as linhas de A na métrica de Φ. Isto é importante, pois mais adiante, em autoestrutura de uma matriz de produto escalar voltaremos a usar esse fato.

iii) Os autovetores M são obtidos como segue:

Da expressão (c8) temos

$$A = ND_\alpha M^t$$

onde:

$N^t\Omega N=I$ e $M^t\Phi M=I$.

então:

$$
\begin{aligned}
A^t &= MD_\alpha^t N^t \\
A^t\Omega &= MD_\alpha^t N^t \Omega \\
A^t\Omega N &= MD_\alpha^t I \\
A^t\Omega N\left(D_\alpha^t\right)^{-1} &= M
\end{aligned}
$$

então:

$$M = A^t\Omega N D_\alpha^{-1} \qquad \text{(c11)}$$

A expressão (c11) é conhecida como nome de: Fórmula de transição (transition formula).

iv) A autoestrutura generalizada na equação (c10) é obtida pelo cálculo da primeira autoestrutura da matriz $\Omega^{1/2}A\Phi A^t(\Omega^{1/2})^t$, isto é:

$$\Omega^{1/2}A\Phi A^t\left(\Omega^{1/2}\right)^t = UD_\lambda U^t \tag{c12}$$

onde os autovetores U têm a normalização usual $U^tU=I$.

Por outro lado, da expressão (c12) obtemos:

$$A\Phi A^t = \Omega^{-1/2}UD_\lambda\left(\Omega^{-1/2}U\right)^t \tag{c13}$$

então, igualando a expressão (c13) com a expressão (c10), temos:

$$N = \Omega^{-1/2}U \qquad D_\alpha = \sqrt{D_\lambda}$$

C.6. Qualidade e Erro de Aproximação.

Sabe-se que o autovalor (valor básico) expressa a magnitude da dimensão de uma matriz. Por outro lado, a norma quadrada de A na expressão (c8) é: $\alpha_1^2+\alpha_2^2+...+\alpha_r^2$. Quando aproximamos A por $\hat{A}_{[k]}$ apenas os k maiores autovalores são considerados, e a norma quadrada de $\hat{A}_{[k]}$ neste caso é: $\alpha_1^2+\alpha_2^2+...+\alpha_k^2$.

Então a **medida de qualidade** de aproximação é:

$$R_{[k]} = \frac{\alpha_1^2+\alpha_2^2+.....+\alpha_k^2}{\alpha_1^2+\alpha_2^2+.....+\alpha_r^2} \tag{c14}$$

Mardia, Kent e Bibby (1979) usa este fato para aproximar uma medida de bondade de ajuste na parte de MDS métrico.

Então da expressão (c14) pode-se notar que: $0 \le R_{[k]} \le 1$. O **erro de aproximação** pode ser medido pelo complementar de $R_{[k]}$ como segue:

$$1-R_{[k]}=\frac{\alpha_{k+1}^2+\alpha_{k+2}^2+.....+\alpha_r^2}{\alpha_1^2+\alpha_2^2+.....\alpha_r^2} \qquad (c15)$$

C.7. Autoestrutura de uma Matriz Produto Escalar.

Vimos que em conseqüência de (ii) e (iii) da DVSG os autovetores esquerdo (basicos) da matriz A são os autovetores de uma matriz simétrica quadrada, ver equação (c10).

Agora vejamos o caso de ter uma matriz de dados, é dizer n indivíduos (ou observações) mensurados em p variáveis expressos na matrizX_{nxp}.

$$X_{nxp}=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}\begin{matrix} \vec{x}_1^t \\ \vec{x}_2^t \\ \vdots \\ \vec{x}_n^t \end{matrix}$$

Seja o **vetor de médias**:

$$\vec{\bar{x}}=\frac{\sum_{i=1}^{n}\vec{x}_i}{n}$$

Seja o vetor:

$$1^t=(1,1,.....,1)_{1xn}$$

Seja a matriz de resíduos:

$$Z=\begin{pmatrix} x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1p}-\bar{x}_p \\ x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2p}-\bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{np}-\bar{x}_p \end{pmatrix}$$

Na forma matricial, Z pode ser escrito da seguinte forma:

$$Z = X - \mathbf{1}\bar{\bar{x}}^t \tag{c16}$$

Então, se temos uma matriz (ZZ^t) isto é:

$$ZZ^t = \left[\left(X - \mathbf{1}\bar{\bar{x}}^t\right)\left(X - \mathbf{1}\bar{\bar{x}}^t\right)^t\right] \tag{c17}$$

Podemos observar que a equação (c17) já é uma matriz simétrica e quadrada, portanto pela autoestrutura básica temos:

$$ZZ^t = \left[\left(X - \mathbf{1}\bar{\bar{x}}^t\right)\left(X - \mathbf{1}\bar{\bar{x}}^t\right)^t\right] = NDmN^t \tag{c18}$$

onde: $N^t N = I$

Então da equação (c18) temos:

$$ZZ^t = ND_\mu^{1/2}D_\mu^{1/2}N^t = ND_\mu^{1/2}\left(ND_\mu^{1/2}\right)^t$$

Portanto, uma aproximação para a matriz Z é :

$$\hat{Z}_{[k]} = N_{[k]}D\mu_{[k]}^{1/2} \tag{c19}$$

Se relacionamos a expressão (c19) às expressãos (c10) então, os valores básicos α_k de Z são a raiz quadrada ($\mu_k^{1/2}$) dos correspondentes autovalores de ZZ^t.

A matriz $C=ZZ^t$ é chamada de **matriz produto escalar de ordem nxn entre as linhas da matriz Z** ou ainda **produto escalar entre as linhas de X com respeito a seu centro de gravidade** (vetor de médias). Nota-se que a partir da matriz produto escalar $C=ZZ^t$ pode-se encontrar uma configuração de pontos em menor dimensão, ajustada para a matriz Z. Na prática isto quer dizer que, se nos não temos a matriz $Z = X - \mathbf{1}\bar{\bar{x}}^t$, mas temos uma matriz de produto escalar $C=ZZ^t$ (entre linhas, neste caso), então a estrutura básica de C fornece uma exposição grafica em menor dimensão como mostra acima a equação (c19).

C.8. Generalizando para um Espaço Euclidiano Ponderado.

Em geral se C é uma matriz produto escalar entre qualquer conjunto de pontos vetoriais e seu centro de gravidade em qualquer espaço euclidiano ou espaço euclidiano ponderado, onde cada ponto é ponderado por seu respectivo elemento da matriz diagonal Dω, então a projeção destes pontos sobre um espaço de menor dimensão (estimado) é obtida através da autoestrutura generalizada (DVSG) de C :

$$C = ND\mu N^t \tag{c20}$$

onde : $N^tD\omega N=I$, (estamos no caso particular em que $\Omega=D\omega$), portanto uma aproximação é:

$$\hat{Z} = N_{[k]}D\mu_{[k]}^{1/2} \tag{c21}$$

Como visto, em conseqüências da DVSG nas expressãos (c12) e (c13), na prática o cálculo consiste de construir a primeira autoestrutura da matriz simétrica $D_\omega^{1/2}CD_\omega^{1/2}$ isto é:

$$D_\omega^{1/2}CD_\omega^{1/2} = UD_\mu U^t \tag{c22}$$

onde os autovetores U têm a normalização usual $U^tU=I$, logo da expressão (c24) temos:

$$C = D_\omega^{-1/2}UD\mu\left(D_\omega^{-1/2}U\right)^t = D_\omega^{-1/2}UD_\mu^{1/2}\left(D_\omega^{-1/2}UD_\mu^{1/2}\right)^t \tag{c23}$$

Então uma aproximação para Z, quando $\Omega=D\omega$, é:

$$F = \hat{Z} = D_\omega^{-1/2}U_{[k]}D\mu_{[k]}^{1/2} \tag{c24}$$

Até aqui vimos que a partir de uma matriz produto escalar no espaço euclidiano ou no espaço euclidiano ponderado, pode estimar uma configuração de pontos através da autoestrutura generalizada (DVSG). Como estamos interessados no procedimento de escalonamento de um conjunto de pontos, vamos trabalhar com as distâncias entre pontos interiores deste conjunto. A notação usual será: os índices i,j indicam elementos de um mesmo conjunto isto é δ_{ij} indica a medida distância entre duas linhas i e j.

Apêndice D

RELAÇÃO ENTRE PRODUTO ESCALAR E DISTÂNCIA

D.1. Definição.

Se temos um conjunto de produtos escalares entre todos os pontos, denotado por **C**:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \qquad \text{(d1)}$$

e temos um conjunto de todas as distâncias entre pontos δ_{ij}^2 (i,j=1,2,...,n), arranjados numa matriz $\Delta^{(2)}$ da seguinte forma:

$$\Delta^{(2)} = \begin{pmatrix} \delta_{11}^2 & \delta_{12}^2 & \cdots & \delta_{1n}^2 \\ \delta_{21}^2 & \delta_{22}^2 & \cdots & \delta_{2n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{n1}^2 & \delta_{n2}^2 & \cdots & \delta_{nn}^2 \end{pmatrix} \qquad \text{(d2)}$$

Então podemos obter uma configuração de pontos para ajustar a matriz $\Delta^{(2)}$, usando o resultado de que a distância quadrada está relacionada com o produto escalar, por:

$$\delta_{ij}^2 = c_{ii} + c_{jj} - 2c_{ij} \qquad \text{(d3)}$$

e além disso pode-se mostrar que a matriz **C**, produto escalar pode ser recuperada a partir da matriz $\Delta^{(2)}$ de distâncias quadradas pelo processo de "Dupla centralização" (Double centering), isto é:

$$C = \theta\left[-\frac{1}{2}\Delta^{(2)}\right]\theta^t \qquad (d4)$$

onde:

θ é uma matriz idempotente de ordem nxn definida da seguinte maneira:

$$\theta = I - \mathbf{1}\left(\mathbf{1}^t w\right)^{-1} w^t \qquad (d5)$$

w : um vetor formado pela diagonal da matriz Dw

$\mathbf{1}$: um vetor de uns isto é: $\mathbf{1}^t = [1,1,...,1]_{1xn}$.

sujeito a restrição: $\mathbf{1}^t w = 1$

Deste modo a matriz **C**, é uma matriz de produto escalar de pontos (ponderados) com respeito a seu centro de gravidade e então as equações (25) e (26) podem ser aplicados para encontrar uma configuração de pontos que reproduza uma estimação das distâncias originais. Todo exposto anteriormente é conhecido com o nome de Escalonamento de Young-Householder (Scaling Young-Householder), chamado muitas vezes por Young e Householder (1938) de Escalonamento Classico (Classical Scaling) ver Krzanowski (1994).

Observação

A relação acima pode ser aplicada a qualquer matriz simétrica de distância quadráticas entre pontos $\Delta^{(2)}$, não necessariamente Euclidiana, para obter uma exposição gráfica em um espaço de menor dimensão. A continuação formalizaremos as expressões d3 e d4 baseados no trabalho de Greenacre (1984).

D.2. Formalização da expressão(d3)

Suponha que temos $\vec{x}_1,\vec{x}_2,...,\vec{x}_n$ pontos vetoriais no espoço p-dimencional arranjados em uma matriz $X=(\vec{x}_1,\vec{x}_2,...,\vec{x}_n)^t$ então a matriz simétrica $\Delta_{nxn}^{(2)}$ de distâncias quadráticas entre os vetores mencionados acima é dado por:

$$\Delta^{(2)} = c_{nx1}\mathbf{1}^t_{1xn} + \mathbf{1}_{nx1}c^t_{1xn} - 2C_{nxn} \tag{d6}$$

onde :
- **C** é uma matriz nxn dos produtos escalares entre os vetores $\vec{x}_1,\vec{x}_2,...,\vec{x}_n$.
- c é um vetor formado pelos elementos da digonal de **C**.
- **1** é um vetor formado por uns.

Prova

Como $\vec{x}_1,\vec{x}_2,...,\vec{x}_n$ são pontos vetoriais no espaço p-dimensional, então tais vetores podem ser representados como uma matriz de dados, isto é:

$$X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}\begin{matrix}\vec{x}_1^t \\ \vec{x}_2^t \\ \vdots \\ \vec{x}_n^t\end{matrix} \tag{d7}$$

Agora como **C** é uma matriz de ordem nxn dos produtos escalares entre os vetores $\vec{x}_1,\vec{x}_2,...,\vec{x}_n$, então a matriz **C** é expresso da seguinte maneira:

$$C=\begin{pmatrix} \vec{x}_1^t\vec{x}_1 & \vec{x}_1^t\vec{x}_2 & \cdots & \vec{x}_1^t\vec{x}_n \\ \vec{x}_2^t\vec{x}_1 & \vec{x}_2^t\vec{x} & \cdots & \vec{x}_2^t\vec{x}_n \\ \vdots & \vdots & \ddots & \cdots \\ \vec{x}_n^t\vec{x}_1 & \vec{x}_n^t\vec{x}_2 & \cdots & \vec{x}_n^t\vec{x}_n \end{pmatrix}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{n2} \end{pmatrix}=X_{nxp}X^t_{pxn}$$

Por outro lado, a distância quadradatica entre os vetores $\vec{x}_i^t$ e $\vec{x}_j^t$ é:

$$\begin{aligned}\delta_{ij}^2 &= \left(\vec{x}_i - \vec{x}_j\right)^t \left(\vec{x}_i - \vec{x}_j\right) \\ &= \vec{x}_i^t \vec{x}_i + \vec{x}_j^t \vec{x}_j - 2\vec{x}_i^t \vec{x}_j \\ &= c_{ii} + c_{jj} - 2c_{ij}\end{aligned}$$

Consequentemente temos que:

$$\delta_{ij}^2 = c_{ii} + c_{jj} - 2c_{ij} \tag{d8}$$

Logo, a equação (d8) é um elemento da equação (d6). Então mostramos que um elemento da matriz $\Delta_{nxn}^{(2)}$ pode ser escrito como a expressão (d8).

D.3. Formalização da expressão (d4).

Agora suponha consequentemente que temos a matriz $\Delta_{nxn}^{(2)}$ de distâncias Euclidianas quadrada entre n pontos vetoriais (cada vetor no espaço p-dimencional), então a matriz **C**, produto escalar relativa para um baricentro de pontos definida pelo vetor de massas w onde: $\mathbf{1}^t\mathbf{w}=1$(restrição importante) é dado por:

$$C = \theta\left(-\frac{1}{2}\Delta^{(2)}\right)\theta^t \tag{d9}$$

onde:

$$\theta = I - \mathbf{1}w^t = I - \mathbf{1}\,(\mathbf{1}^t w)^{-1} w^t \tag{d10}$$

Prova

A partir da matriz X especificada na equação (d7) podemos construir a matriz da equação (d6) isto é: $\Delta^{(2)} = c\mathbf{1}^t + \mathbf{1}c^t - 2C$, lembrando que temos uma restrição importante: $\mathbf{1}^t w = w^t \mathbf{1} = 1$.

A partir da equação (d10) temos:

$$\begin{aligned}\theta \mathbf{1}\mathbf{c}^t &= \left(\mathrm{I} - \mathbf{1}\mathbf{w}^t\right)\mathbf{1}\mathbf{c}^t \\ &= \mathbf{1}\mathbf{c}^t - \mathbf{1}\mathbf{w}^t\mathbf{1}\mathbf{c}^t \\ &= \mathbf{1}\mathbf{c}^t - \mathbf{1}\mathbf{c}^t \\ &= 0\end{aligned}$$

então:

$$\theta \mathbf{1}\mathbf{c}^t = 0 \qquad \text{(d11)}$$

Por outro lado, temos que:

$$\begin{aligned}\mathbf{c}\mathbf{1}^t\theta^t &= \left[\left(\mathrm{I} - \mathbf{1}\mathbf{w}^t\right)\mathbf{1}\mathbf{c}^t\right]^t \\ &= \left[\mathbf{1}\mathbf{c}^t - \mathbf{1}\mathbf{w}^t\mathbf{1}\mathbf{c}^t\right]^t \\ &= \mathbf{c}\mathbf{1}^t - \mathbf{c}\mathbf{1}^t\mathbf{w}\mathbf{1}^t \\ &= \mathbf{c}\mathbf{1}^t - \mathbf{c}\mathbf{1}^t \\ &= 0\end{aligned}$$

então:

$$\mathbf{c}\,\mathbf{1}^t\theta^t = 0 \qquad \text{(d12)}$$

Agora, substituindo as expressões (d11) e (d12) na expressão (d6) temos:

$$\begin{aligned}\theta\Delta^{(2)}\theta^t &= \theta\left(\mathbf{c}\mathbf{1}^t + \mathbf{1}\mathbf{c}^t - 2\mathrm{C}\right)\theta^t \\ &= \theta\mathbf{c}\mathbf{1}^t\theta^t + \theta\mathbf{1}\mathbf{c}^t\theta^t - 2\theta\mathrm{C}\theta^t \\ &= -2\theta\mathrm{C}\theta^t\end{aligned}$$

Então mostramos que:

$$\theta\Delta^{(2)}\theta^t = -2\theta\mathrm{C}\theta^t \qquad \text{(d13)}$$

Agora, falta mostrar que:

$$\theta C \theta^t = C \tag{d14}$$

Temos:

$$\begin{aligned} \theta C \theta^t &= \left(I - 1w^t\right) C \left(I - 1w^t\right)^t \\ &= C - 1w^t C - Cw1^t + 1w^t Cw1^t \end{aligned} \tag{d15}$$

e para mostrar que a expressão (d14) é valida basta mostrar que $Cw=0$ na expressão (d15).

Por outro lado, como **C** é uma matriz produto escalar relativo a um baricentro (que poderia ser o vetor de médias) então a matriz X, como exemplificada na equação (d7), seria da seguinte maneira:

$$Z = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \cdots & x_{np} - \bar{x}_p \end{pmatrix} \begin{matrix} \vec{x}_1^t \\ \vec{x}_2^t \\ \vdots \\ \vec{x}_n^t \end{matrix}$$

Então a matriz de produto escalar **C** fica da seguinte forma:

$$C = \begin{pmatrix} \left(\vec{x}_1 - \vec{\bar{x}}\right)^t \left(\vec{x}_1 - \vec{\bar{x}}\right) & \left(\vec{x}_1 - \vec{\bar{x}}\right)^t \left(\vec{x}_2 - \vec{\bar{x}}\right) & \cdots & \left(\vec{x}_1 - \vec{\bar{x}}\right)^t \left(\vec{x}_n - \vec{\bar{x}}\right) \\ \left(\vec{x}_2 - \vec{\bar{x}}\right)^t \left(\vec{x}_1 - \vec{\bar{x}}\right) & \left(\vec{x}_2 - \vec{\bar{x}}\right)^t \left(\vec{x}_2 - \vec{\bar{x}}\right) & \cdots & \left(\vec{x}_2 - \vec{\bar{x}}\right)^t \left(\vec{x}_n - \vec{\bar{x}}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left(\vec{x}_1 - \vec{\bar{x}}\right) & \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left(\vec{x}_2 - \vec{\bar{x}}\right) & \cdots & \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left(\vec{x}_n - \vec{\bar{x}}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\vec{x}_n - \vec{\bar{x}}\right)^t \left(\vec{x}_1 - \vec{\bar{x}}\right) & \left(\vec{x}_n - \vec{\bar{x}}\right)^t \left(\vec{x}_2 - \vec{\bar{x}}\right) & \cdots & \left(\vec{x}_n - \vec{\bar{x}}\right)^t \left(\vec{x}_n - \vec{\bar{x}}\right) \end{pmatrix}_{nxn}$$

Sabe-se que : $w^t=[w_1, w_2,..., w_n]$ e que $1^t w=1$ isto implica que $\sum_{i=1}^{n} \omega_i = 1$.

$C_{nxn}\, w_{nx1}$ é um vetor; o j-esimo elemento do vetor Cw é $(Cw)_j$ o qual é expresso da seguinte maneira, para todo j=1,2,...,n :

$$
\begin{aligned}
(Cw)_j &= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left(\vec{x}_1 - \vec{\bar{x}}\right) w_1 + \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left(\vec{x}_2 - \vec{\bar{x}}\right) w_2 + \ldots \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left(\vec{x}_n - \vec{\bar{x}}\right) w_n \\
&= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \sum_{i=1}^{n} \left(\vec{x}_i - \vec{\bar{x}}\right) w_i \\
&= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left[\sum_{i=1}^{n} \vec{x}_i w_i - \sum_{i=1}^{n} \vec{\bar{x}} w_i \right] \\
&= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left[\sum_{i=1}^{n} \vec{x}_i w_i - \vec{\bar{x}} \sum_{i=1}^{n} w_i \right] \\
&= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left[\sum_{i=1}^{n} \vec{x}_i w_i - \vec{\bar{x}} \right] \\
&\text{agora tomando : } w_i = \frac{1}{n} \\
&= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left[\frac{1}{n} \sum_{i=1}^{n} \vec{x}_i - \vec{\bar{x}} \right] \\
&= \left(\vec{x}_j - \vec{\bar{x}}\right)^t \left[\vec{\bar{x}} - \vec{\bar{x}} \right] \\
&= 0
\end{aligned}
$$

Então mostramos que o j-ésimo elemento do vetor Cw é zero isto é: $(Cw)_j=0$, para todo j=1,2,...,n. Logo a expressão (d14) fica provado, implicando que a expressão (d13) é dada por:

$$\theta \Delta^{(2)} \theta^t = -2C$$

isto é, está provado que:

$$C = \theta\left(-\frac{1}{2}\Delta^{(2)}\right)\theta^t$$

Observações

1.- Quando $w^t=[1/n, 1/n,..., 1/n]$ temos que $\theta=\mathbf{H}$ (matriz idempotente) onde H é conhecida como matiz de projeção, isto é:

$$H = \left[I - \frac{1}{n} \mathbf{1}\mathbf{1}^t \right]$$

Greenacre (1984) afirma que também pode ser provado para um caso geral.

A observação 1 será usada foi usada no capítulo II.

2.- A matriz X foi trasladada a origem, para facilitar a formalização da expressão (d4).

3.- O resultado da expressão (d4) é independente da origem isto é $s_{ij} = \left(\vec{x}_i - \vec{a}\right)^t \left(\vec{x}_i - \vec{a}\right)$. Isto quer dizer o vetor "a" não necessariamente tem que ser o vetor de médias.

4.- A transformação $\theta(-1/2\Delta^{(2)})\theta^t$ é chamada de Dupla Centralização.

Apêndice E

REGRESSÃO MONOTÔNICA OU ISOTÔNICA

E.1. Definição matemática formal de uma funcão Isotônica.

Dado X, um conjunto finito $\{x_1, x_2, \ldots, x_k\}$, ordenado da seguinte maneira $x_1 \leq x_2 \leq \ldots \leq x_k$.

Uma função f sob X é isotônico (com respeito a essa ordem) se: $f(x_1) \leq f(x_2) \leq \ldots \leq f(x_k)$.

E.2. Definição matemática formal de uma Regressão Isotônica.

Suponha g uma função dada sobre X. A função g^* sobre X é uma Regressão Isotônica de g com pesos w, se e somente se, g^* é isotônica e minimiza:

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x) \qquad \text{(e1)}$$

na clases de todas as funções isotônicas de f sob X.

Maiores detalhes sobre Regressão Isotônica ver Barlow, Bartholomew, Bremner e Brunk (1972).

E.3. No contexto de Escalonamento Multidimensional

Um conjunto formado de **n** similaridades ou distâncias observadas δ_{ij} são ordenadas tal que:

$$\delta_{i_1 j_1} \leq \delta_{i_2 j_2} \leq \ldots\ldots\ldots\ldots \leq \delta_{i_n j_n}$$

Por outro lado a partir das **n** similaridades ou distâncias pode-se achar uma configuração no espaço p-dimensional, no qual vamos ter um conjunto de distâncias associadas d_{ij}, isto é:

$$d_{i_1 j_1} \leq d_{i_2 j_2} \leq \ldots\ldots\ldots\ldots \leq d_{i_n j_n}$$

e além disso consideremos um conjunto de pesos da seguinte forma:

$$w_{i_1 j_1} \leq w_{i_2 j_2} \leq \ldots\ldots\ldots\ldots \leq w_{i_n j_n}$$

Para não haver confusão adotamos uma notação simplificada assim:

$$\delta_{i_k j_k} = \delta_k \qquad d_{i_k j_k} = d_k \qquad w_{i_k j_k} = w_k$$

E.3.1. Definição de uma Regressão Isotônica no contexto de Escalonamento Multidimensional

Uma Regressão Monotônica ou Isotônica de d_i sobre um conjunto ordenado δ_i com pesos w_i é dado por uma função $f(d_i)$ que minimiza:

$$\sum_{i=1}^{m} \left[d_i - f(d_i) \right]^2 w_i \qquad \text{(e2)}$$

sujeito a restrição de monotocidade, isto é,

$$i \leq j \Rightarrow f(d_i) \leq f(d_j)$$

Observação

Nota-se que a função **g(x)** que minimiza a equação (e1) é a função identidade na equação (e2), isto é, $g(d_i)=d_i$, e ao contrário do que acontece na regressão linear simples não ha necessidade de se fazer suposições sobre a forma analítica da função g(x). Se $w_i=1$ para todo i=1,2,....,n então a Regressão monotônica utiliza um método de quadrados mínimos. Se $w_i \neq 1$ para todo i=1,2,....,n então a Regressão monotônica utiliza o método de quadrados mínimos ponderados, que são resolvidos por métodos iterativos. Barlow(1972) prova que a Regressão monotônica existe e é única.

Existem algoritmos para estimar a função $f(d_i)$, como por exemplo :

- "**POOL-ADJACENT-VIOLATORS ALGORITM**"
- "**UP-AND-DOWN-BLOCKS ALGORITM** ".

Kruskal (1964) usa o algoritmo "Up-and-down-blocks" no algoritmo "steep decente" para a parte MDS não métrico.

A seguir mostramos como trabalha o Algoritmo Pool-adyacent-violators.

E.4. Algoritmo Pool-Adyacent-Violators.

Como dados iniciais temos:
δ_i : Variável regressora ou independente.
$\mathbf{d_i}$: Variável resposta ou dependente.
$\mathbf{w_i}$: Peso.
$\mathbf{n_i}$: Define a estrutura bloco.

Passo 1

- Ordenar os δ_i (ordem crescente) e logicamente vai mudar d_i .
- Fixar : $\hat{\mathbf{d}}_i = \mathbf{d_i}$, $\hat{\mathbf{w}}_i = \mathbf{w_i}$ e $\mathbf{n_i = 1}$. i=1,2,...,n.

Passo 2

Se: $\hat{d}_1 \leq \hat{d}_2 \leq \ldots\ldots\ldots \leq \hat{d}_n$ então finaliza, caso contrário passa para o passo 3.

Passo 3

Selecionar qualquer par de blocos que "**viole**" o ordenamento. Formalmente se bloco **i** e **j** **(i<j)** viola o ordenamento então temos que o novo $\hat{d}_i$ estimado é dado por:

$$\hat{d}_i = \frac{\hat{w}_i \hat{d}_i + \hat{w}_j \hat{d}_j}{\hat{w}_i + \hat{w}_j}$$

e então fixamos:

$\hat{d}_k = \hat{d}_i$ **k=i+1,...,i+n_i+n_j-1** $\hat{w}_i = \hat{w}_i + \hat{w}_j$ **w_j=0**

$n_i = n_i + n_j$ **n_j=0**

Logo retornar para o **Passo 2**.

A seguir mostramos um exemplo prático de como o algoritmo funciona.

i) Primeiro mostramos os dados na tabela 1. Pode-se observar que os dados já foram ordenados de forma crescente.

TABELA 1

i	δ_i	$d_i=\hat{d}_i$	$w_i=\hat{w}_i$	n_i
1	0.8	0.25	1.00	1
2	1.2	2.00	0.25	1
3	1.5	1.00	0.75	1
4	2.5	0.50	1.50	1
5	3.0	2.00	1.00	1

ii) Agora buscamos o bloco que viola o ordenamento.

Claramente notamos que em (i=2 e j=3) viola o ordenamento. Logo, pelo passo 3 temos:

$$\hat{d}_{i=2} = \frac{\hat{w}_{i=2}\hat{d}_{i=2} + \hat{w}_{j=3}\hat{d}_{j=3}}{\hat{w}_{i=2} + \hat{w}_{j=3}} = \frac{(0.25)(2)+(0.75)(1)}{0.25+0.75} = 1.25$$

para: i=2 e j=3 temos n_2=1 e n_3=1

Agora substituindo em k temos: k=3.

onde :

$$\hat{d}_{k=3} = \hat{d}_{i=2} = 1.25$$

$$\hat{w}_{i=2} = \hat{w}_{i=2} + \hat{w}_{j=3} = 0.25 + 0.75 = 1 \quad w_{j=3}=0$$

$n_{i=2} = n_{i=2} + n_{j=3} = 1+1=2$ $\quad n_3=0$

iii) A seguir construímos novamente uma tabela (tabela 2) com os dados obtidos em (ii).

TABELA 2

i	δ_i	$\mathbf{d_i = \hat{d}_i}$	$\mathbf{w_i = \hat{w}_i}$	$\mathbf{n_i}$
1	0.8	0.25	1.00	1
2	1.2	1.25	1.00	2
3	1.5	1.25	0.00	0
4	2.5	0.50	1.50	1
5	3.0	2.00	1.00	1

iv) Novamente procuramos o bloco que viola o ordenamento. Percebe-se que o ordenamento é violado no bloco (i=2 e j=4). Agora pelo passo 3 temos:

$$\hat{d}_{i=2} = \frac{\hat{w}_{i=2}\hat{d}_{i=2} + \hat{w}_{j=4}\hat{d}_{j=4}}{\hat{w}_{i=2} + \hat{w}_{j=4}} = \frac{(1)(1.25)+(1.50)(0.5)}{1.50+1} = 0.8$$

para: i=2 e j=4 temos n_2=2 e n_4=1

Agora substituindo em k temos: k=3,4.

então :

$$\hat{d}_{k=3} = \hat{d}_{k=4} = \hat{d}_{i=2} = 0.8$$

$$\hat{w}_{i=2} = \hat{w}_{i=2} + \hat{w}_{j=4} = 1+1.5 = 2.5 \qquad w_{j=4}=0$$

$n_{i=2} = n_{i=2} + n_{j=4} = 1+1=2$ $\quad n_4=0$

v) Agora arranjando os dados obtidos em (iv) na tabela 3 temos:

TABELA 3

i	δ_i	$\mathbf{d_i}=\hat{d}_i$	$\mathbf{w_i}=\hat{w}_i$	$\mathbf{n_i}$
1	0.8	0.25	1.0	1
2	1.2	0.8	2.5	3
3	1.5	0.8	0.0	0
4	2.5	0.8	0.0	0
5	3.0	2.0	1.0	1

vi) Logo na tabela 3 encontra-se a transformação monotônica de d_i em δ_i.

Apêndice F

NÍVEIS DE MENSURAÇÃO OU MEDIDA

F.1. Definição de Medida

Conceitualmente, medida pode ser definida como uma maneira de obter símbolos (ou características) para representar propriedades de indivíduos, objetos, eventos, condições ou situações estas, nas quais tais símbolos tenham parantesco.

Existem 4 tipos comuns de medições, nominal, ordinal, intervalar e razão, os quais serão detalhados a seguir baseados nos tabalhos de Green e Tull (1978), Schiffman, Reynolds e Young (1981) e Cambillo (1996).

F.1.1. Escala Nominal

Neste tipo de escala, os números servem só como etiquetas ou chapas para identificar objetos, propriedades ou eventos. Associam-se os objetos a categorias (ou conjuntos) mutuamente exclusivos.

A cada conjunto atribui-se um número, por exemplo a variável sexo pode pegar dois valores femenino=1, masculino=2. Estes números são apenas identificadores, pois não existe uma relação de ordem entre elas. O número só indica a inclusão de uma determinada classe. Se os objetos têm o mesmo número, então pertencem à mesma classe e pode-se dizer que são equivalentes. Este tipo de variáveis é freqüentemente usado em pesquisa de mercado: sexo, zona geográfica, profissão, estado civil, etc.

Claramente, as operações que podemos realizar com as variáveis nominais estão baseadas nas relações de equivalência, só são permitidas operações matemáticas elementares. Na parte da estatística, podem-se fazer alguns testes em tabelas de contingência, freqüências,

modas, etc. ainda que as usuais operações estatísticas como média, desvio padrão, etc. não sejam empiricamente importantes ou significativas.

F.1.2. Escala Ordinal

Escala ordinal, como o nome sugere, é uma escala relacionada à ordem ou escala de posto. É obtida através da hierarquização dos objetos, colocando-os numa ordem relacionada com o grau na que possuem a característica. Utilizam-se muito em pesquisa de mercados, dentro dos estudos para codificar as preferências dos consumidores.

Por exemplo, se um entrevistado ordena 6 marcas (A, B, C, D e E) de bebida segundo suas preferências, obteremos por exemplo, A1, B2, E3, F4, C5, D6. Isto não significa que a marca D goste o dobro que E, nem sequer que a diferença entre A e B seja a mesma que E y F. Simplesmente indica um ordem de preferência.

Uma escala ordinal posiciona toda a informação de uma escala nominal. Por outro lado em um delineamento com escalas ordinais, descrições estatísticas podem empregar medidas posicionais como mediana, quartil, percentil, ou outras estatísticas de acordo com o ordem, como fazer testes não paramétricos. Nota-se também que a média aritmética usual não pode ser significativamente interpretada com os postos dos dados.

F.1.3. Escala Intervalar

Neste tipo de escalas não existe um zero natural, no entanto a distância entre dois números segue tendo um significado como ocorre com a temperatura. Os exemplos mais comuns de escala intervalar são as escalas Farenheit(F) e Centígrado(C) usadas para medir temperatura. O ponto zero destas escalas é arbitrário, isto é, não podemos dizer que o estado de qualquer valor sobre uma escala intervalar específico é um múltiplo de outro. Agora, vejamos um exemplo para esclarecer este ponto.

Não é empiricamente correto dizer que um objeto com uma temperatura de 50^0 F é duas vezes mas quente que outro objeto com temperatura de 25^0 F. Lembrando que a fórmula de conversão de Farenheit para centígrado é dada por:

$$T_C = \frac{5}{9}\left(T_F - 32\right)$$

fazendo os cálculos respectivos de conversão temos: que 50^0F eqüivale a 10^0C e 25^0F eqüivale a -3.9^0C . Percebemos que os graus centígrados não são a razão de 2:1. Entretanto, podemos afirmar que diferença entre valores sobre diferentes escalas de temperatura são múltiplos. Por exemplo, a diferença de (50^0F-0^0F) é duas vezes a diferença de (25^0F-0^0F), da mesma maneira usando a formula de conversão para escala centigrado temos que: [10^0C-(-17.7^0C)]=27.7^0C e [3.9^0C-(-17.7^0C)]=13.8^0C, onde percebemos que a ordem da razão 2:1 é mantida nos graus centígrados.

Na escala intervalar podem ser usadas transformações lineares da forma: $y=a+bx$; $b>0$. Médias de escala intervalar podem ser transformadas para outras médias através de uma transformação linear positiva.

A maior parte das estatísticas ordinárias como a média aritmética, desvio padrão, coeficiente de correlação, testes paramétricos, etc. requer apenas de escalas intervalares para podermos fazer os cálculos e termos interpretações significativamente importantes. Não obstante algumas estatísticas como a média geométrica não, são empiricamente significativas se aplicadas para dados de escala intervalar.

Esta escala é usada com freqüência na pesquisa de mercados para medir as atitudes, como, por exemplo, uma escala de desejo de compra com dez categorias (1-10). Assim, pode-se dizer que pessoas com notas respectivas 4 e 6 diferenciam-se em desejo de compra no mesmo grau que duas pessoas com notas 8 e 10, no entanto a pessoa com nota 8 não tem o duplo desejo de compra que a pessoa de nota 4.

F.1.4. ESCALA RAZÃO

É a escala métrica mais perfeita. Tem sua origem no zero com um sentido real e invariável, por exemplo, o zero em ingressos significa uma ausência total de ingressos, os intervalos entre os números tem um significado e podem-se fazer comparações.

A escala razão representa a elite das escalas, onde todas as operações matemáticas clássicas e as operações estatísticas são permissíveis. Estas escalas são usualmente fundadas nas ciências fisicas (isto é escalas para medir comprimento, peso, etc.). São as escalas que tem as maiores propriedades, no entanto estas escalas são as menos usadas em pesquisa de mercado.

Escala razão também pode ser definida da seguinte maneira; objetos são colocados sobre uma escala tal que a posição através da escala representa a magnitude absoluta dos atributos.

Os quatro tipos de escalas, estão ligados por uma relação de hierarquia (de menos importante a mais importante), isto é desde, a escala nominal que é a menos operativa, até a escala razão que tem as maiores propriedades.

As variáveis relacionadas com a atitude, preferência, intenções, avaliações de atributos, etc, podem ser medidas nas escalas anteriores. É óbvio que uma escala de razão ou intervalar possua maiores propriedades do que a de ordinal ou nominal, no entanto resulta difícil de obter na pratica. As escalas que se usa com freqüência, é a escala ordinal que possui propriedades intermédiarias.

Apêndice G

PROGRAMAS UTILIZADOS ATRAVÉS DO SOFTWARE SAS

H.1. Desenvolvimento dos programas utilizados.

Os dados originais cedidos pela empresa envolvida no projeto encontrava-se no arquivo de nome A. A seguir mostramos a base de dados:

```
PROC PRINT DATA=A;
  RUN;
```

OBS	DIRETOR	E1	E2	E3	E4	E5	E6	E7
1	2	3	4	5	7	1	2	6
2	4	7	6	1	5	3	2	4
3	1	7	1	2	5	6	4	3
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
511	3	7	4	6	5	3	1	2
512	4	5	7	4	6	2	3	1
513	3	7	5	4	6	2	1	3

PROGRAMA 1

Este programa foi feito para dividir a base de dados A em 12 novas, onde cada uma delas correspondia a uma diretoria.

```
OPTIONS OBS=MAX;
DATA DIR1 DIR2 DIR3 DIR4 DIR5 DIR6 DIR7 DIR8 DIR9 DIR10 DIR11 DIR12;
 SET A;
  KEEP F1-F7;
      IF DIRETOR=1 THEN OUTPUT DIR1;
          ELSE IF DIRETOR=2 THEN OUTPUT DIR2;
          ELSE IF DIRETOR=3 THEN OUTPUT DIR3;
          ELSE IF DIRETOR=4 THEN OUTPUT DIR4;
          ELSE IF DIRETOR=5 THEN OUTPUT DIR5;
          ELSE IF DIRETOR=6 THEN OUTPUT DIR6;
          ELSE IF DIRETOR=7 THEN OUTPUT DIR7;
          ELSE IF DIRETOR=8 THEN OUTPUT DIR8;
          ELSE IF DIRETOR=9 THEN OUTPUT DIR9;
          ELSE IF DIRETOR=10 THEN OUTPUT DIR10;
          ELSE IF DIRETOR=11 THEN OUTPUT DIR11;
          ELSE IF DIRETOR=12 THEN OUTPUT DIR12;
 RUN;
```

Por exemplo, o conteúdo da diretoria 1 foi:

OBS	E1	E2	E3	E4	E5	E6	E7
1	7	1	2	5	6	4	3
2	7	5	4	6	1	3	2
3	4	6	1	7	3	5	2
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
47	7	4	1	6	5	2	3
48	7	1	2	3	4	5	6
49	7	4	6	5	2	3	1

PROGRAMA 2

Este programa foi elaborado com base no procedimento de frequências "PROC FREQ" do SAS. Este procedimento foi feito para cada diretoria separadamente, por exemplo, para diretoria 1 temos:

```
PROC FREQ DATA=DIR1;
  RUN;
```

O resultado obtido pelo programa 2 foi arranjado em uma tabela de contingência da seguinte maneira:

RPTA	E1	E2	E3	E4	E5	E6	E7
R1	3	4	20	0	2	6	14
R2	3	2	11	1	7	10	15
R3	3	7	5	6	13	9	6
R4	2	8	8	3	11	8	9
R5	7	7	2	13	8	10	2
R6	8	9	3	16	8	3	2
R7	23	12	0	10	0	3	1

PROGRAMA 3

Este programa foi elaborado com base no procedimento de linguagem matricial interativo "PROC IML" do SAS. O programa foi feito para calcular a distância Qui-quadrada entre os estímulos de cada diretoria separadamente, por exemplo, para diretoria 1 temos:

```
PROC IML;
Y = {3 4 20 0 2 6 14,
     3 2 11 1 7 10 15,
```

```
    3 7 5 6 13 9 6,
    2 8 8 3 11 8 9,        /* Y contém as frequências da diretoria 1 */
    7 7 2 13 8 10 2,
    8 9 3 16 8 3 2,
    23 12 0 10 0 3 1};
X1=(Y)`;              /* X1: é a matriz transposta de Y */
 TOTAL=sum(X1);  /* TOTAL: está indicando a soma de todos os elementos da matriz X1 */
        /* Programa para construir a matriz diagonal D (ver apêndice A, matriz a31) */
c=j(7,1,.);
   n=7;            /* Número de linhas da matriz X1 */
     do i=1 to 7;
         c[i]=sum(X1[,i]);
     end;
pond=c/total;
D=diag(pond);
X=X1/49;      /* 49 indica o total de indivíduos da diretoria 1 */
                  /* Este programa vai calcular as distâncias Qui-quadradas */
 dist=j(7,7,.);
      n=7;      /* n indica o número de estímulos */
        do j=1 to n-1;
           do i=1 to n-j;
              dist[i+j,i]=(X[i,]-X[i+j,])*INV(D)*(X[i,]-X[i+j,])`;
           end;
        end;
  do i=1 to 7;
     dist[i,i]=0;
  end;
* print dist;
/* Agora todas as distâncias calculadas serão arrumadas em uma matriz */
 Create MATDIS from dist;
   append from dist;
```

```
    quit;
PROC PRINT DATA=MATDIS;
  RUN;
```

O resultado obtido pelo programa 3 foi o seguinte:

OBS	COL1	COL2	COL3	COL4	COL5	COL6	COL7
1	0.00000	.	.	.	.	.	.
2	0.51312	0.00000	.	.	.	.	.
3	2.83382	1.59184	0.00000	.	.	.	.
4	0.85131	0.38484	2.67055	0.00000	.	.	.
5	2.12245	0.64140	1.38192	0.99708	0.00000	.	.
6	1.64431	0.57726	0.83382	1.10204	0.25656	0.00000	.
7	2.53061	1.35860	0.16327	2.40816	0.97376	0.48980	0

PROGRAMA 4

Este programa foi elaborado com base no procedimento de Escalonamento Multidimensional "PROC MDS" do SAS. O programa foi feito para cada diretoria separadamente, por exemplo, para diretoria 1 temos:

```
DATA MDSDIR1;
INPUT (EST1 EST2 EST3 EST4 EST5 EST6 EST7) (8.) @54 ESTÍMULOS $6.;
CARDS;
0.0000                                                    EST1
0.5130  0.0000                                            EST2
2.8330  1.5914  0.0000                                   EST3
0.8511  0.3847  2.6698  0.0000                           EST4
2.1218  0.6412  1.3815  0.9968  0.0000                   EST5
1.6438  0.5771  0.8336  1.1017  0.2565  0.0000           EST6
2.5299  1.3582  0.1632  2.4074  0.9735  0.4896  0.0000   EST7
;

PROC MDS DATA=MDSDIR1
  LEVEL=ORDINAL
```

```
    OUT=OUT1 OUTRES=OUT2  OUTFIT=OUT3
      PCOEF PCONFIG PDATA
          FIT=1  PFINAL  PFIT  PFITROW
              PINAVDATA PINEIGVAL
                  PININ PINIT PITER;
  RUN;
PROC PRINT DATA=OUT1;
  RUN;
PROC PRINT DATA=OUT2;
  RUN;
PROC PRINT DATA=OUT3;
  RUN;
```

O resultado do programa final (resumido) foi o seguinte:

MATRIX	Number of Nonmissing Data	Weight	Badness-of-Fit Criterion	Distance Correlation	Uncorrected Distance Correlation
1	21	1	0.00	1.00	1.00

OBS	_DIMENS_	_MATRIX_	_TYPE_	_LABEL_	_NAME_	DIM1	DIM2
1	2	.		CRITERION (STRESS)		0.00390	
2	2	.	CONFIG	EST1	EST1	1.73514	0.92212
3	2	.	CONFIG	EST2	EST2	0.68235	0.23358
4	2	.	CONFIG	EST3	EST3	-1.79318	0.46527
5	2	.	CONFIG	EST4	EST4	1.52092	-0.60518
6	2	.	CONFIG	EST5	EST5	-0.25305	-0.98792
7	2	.	CONFIG	EST6	EST6	-0.43274	-0.28762
8	2	.	CONFIG	EST7	EST7	-1.45943	0.25976

PROGRAMA 5

Este programa foi elaborado com base no procedimento da Análise de Correspondência "PROC CORRESP" do SAS. O programa foi feito para cada diretoria separadamente, por exemplo, para diretoria 1 temos:

```
OPTIONS NODATE PS=60;
TITLE 'GRUPO 1 - 49 INDIVIDUOS' ;
DATA DATA1;
    INPUT RESPOSTA$ ESTIMUL1 ESTIMUL2 ESTIMUL3 ESTIMUL4
          ESTIMUL5 ESTIMUL6 ESTIMUL7;
    LABEL ESTIMUL1='EST1' ESTIMUL2='EST2' ESTIMUL3='EST3'
          ESTIMUL4='EST4' ESTIMUL5='EST5' ESTIMUL6='EST6'
          ESTIMUL7='EST7';
CARDS;
R1  3   4   20   0   2   6   14
R2  3   2   11   1   7   10  15
R3  3   7   5    6   13  9   6
R4  2   8   8    3   11  8   9
R5  7   7   2    13  8   10  2
R6  8   9   3    16  8   3   2
R7  23  12  0    10  0   3   1
;
PROC CORRESP DATA=DATA1 OUT=DATA31 OBSERVED CP RP CELLCHI2
                                   SHORT PROFILE=BOTH ;
    VAR ESTIMUL1 ESTIMUL2 ESTIMUL3 ESTIMUL4 ESTIMUL5 ESTIMUL6
        ESTIMUL7;
    ID RESPOSTA;

PROC PRINT DATA=DATA31;
RUN;
```

PROGRAMA 6

Este programa foi elaborado com base no procedimento de Escalonamento Multidimensional "PROC MDS" do SAS. A diferença com o programa 4 é que agora vamos

trabalhar conjuntamente com as 12 diretorias (este tipo de análise é considerado como MDS sobre estrutura de grupos). O programa foi o seguinte:

```
DATA PONDER;
INPUT (EST1 EST2 EST3 EST4 EST5 EST6 EST7)(7.) @49 ESTIMULO $6.;
CARDS;
0.0000                                                    EST1
0.5130  0.0000                                            EST2
2.8330  1.5914  0.0000                                   EST3
0.8511  0.3847  2.6698  0.0000                           EST4
2.1218  0.6412  1.3815  0.9968  0.0000                   EST5
1.6438  0.5771  0.8336  1.1017  0.2565  0.0000           EST6
2.5299  1.3582  0.1632  2.4074  0.9735  0.4896  0.0000   EST7
0.0000                                                    EST1
0.1865  0.0000                                            EST2
2.2151  3.0719  0.0000                                   EST3
1.0492  1.7487  3.1652  0.0000                           EST4
1.3640  2.0519  1.1950  2.3433  0.0000                   EST5
1.5564  2.3549  1.3757  2.7922  0.5887  0.0000           EST6
1.7021  2.4949  0.2332  2.4832  0.5713  0.8336  0.0000   EST7

   /* E assim sucesivamente coloca-se as matrizes de distâncias para as demais diretorias */

0.0000                                                    EST1
1.6563  0.0000                                            EST2
4.0579  3.0642  0.0000                                   EST3
3.3126  0.7453  2.6501  0.0000                           EST4
3.5611  2.4016  0.8282  2.0704  0.0000                   EST5
4.1408  3.7267  1.4907  3.3954  1.4079  0.0000           EST6
5.0517  3.8095  0.9938  3.8095  2.2360  2.9814  0.0000   EST7
;
```

```
PROC MDS DATA=UNIBPOND
        CONDITION=ROW
          LEVEL=ORDINAL
           COEF=DIAGONAL
            DIMENSION=2
              PFINAL
                OUT=OUT1 OUTRES=OUT2 OUTFIT=OUT3
                   PCOEF PCONFIG PDATA  PFIT PFITROW PINAVDATA
                     PINEIGVAL  PININ PINIT PITER;
  RUN;
PROC PRINT DATA=OUT1;
  RUN;
PROC PRINT DATA=OUT2;
  RUN;
PROC PRINT DATA=OUT3;
  RUN;
```

O resultado do programa final (resumido) foi a seguinte:

OBS	_DIMENS_	_MATRIX_	_TYPE_	_LABEL_	_NAME_	DIM1	DIM2
1	2	.	CRITERION (S-STRESS)			0.06432	
2	2	.	CONFIG	EST1	EST1	0.98854	1.20943
3	2	.	CONFIG	EST2	EST2	1.10307	0.43797
4	2	.	CONFIG	EST3	EST3	-1.36786	1.29185
5	2	.	CONFIG	EST4	EST4	1.13100	-0.72014
6	2	.	CONFIG	EST5	EST5	-0.45422	-1.65857
7	2	.	CONFIG	EST6	EST6	-0.21614	0.07331
8	2	.	CONFIG	EST7	EST7	-1.18439	-0.63384
9	2	1	DIAGCOEF (pesos)			1.26876	0.62471
10	2	2	DIAGCOEF (pesos)			1.34255	0.44448
11	2	3	DIAGCOEF (pesos)			1.02253	0.97695
12	2	4	DIAGCOEF (pesos)			1.27502	0.61183
13	2	5	DIAGCOEF (pesos)			1.27842	0.60469
14	2	6	DIAGCOEF (pesos)			1.12971	0.85074
15	2	7	DIAGCOEF (pesos)			1.23431	0.69028
16	2	8	DIAGCOEF (pesos)			0.99916	1.00084
17	2	9	DIAGCOEF (pesos)			1.27349	0.61500
18	2	10	DIAGCOEF (pesos)			1.28174	0.59762
19	2	11	DIAGCOEF (pesos)			1.09124	0.89955
20	2	12	DIAGCOEF (pesos)			1.10363	0.88431

PROGRAMA 7

Este programa foi elaborado com base no procedimento para encontrar transformações de variáveis lineares e não lineares usando o método de quadrados mínimos alternados para otimizar o ajuste dos dados de uma Regressão linear, "PROC TRANSREG" do SAS. O programa foi feito para cada diretoria separadamente, por exemplo, o programa para diretoria 1 foi:

```
/* UMA FORMA DE JUNTAR ARQUIVOS E ADICIONAR VARIÁVEIS */

DATA MODDIR1;
SET DIR1;
PUT IND1-IND49;
INPUT  DIM1  DIM2  SOMA  DIM12  DIM22  DIM1DIM2;
CARDS;
 1.73514   0.92212   3.86102   3.01071   0.85030   1.60000
 0.68235   0.23358   0.52016   0.46560   0.05455   0.15938
-1.79318   0.46527   3.43197   3.21549   0.21647  -0.83431
 1.52092  -0.60518   2.67794   2.31319   0.36624  -0.92043
-0.25305  -0.98792   1.04002   0.06403   0.97598   0.24999
-0.43274  -0.28762   0.26998   0.18726   0.08272   0.12446
-1.45943   0.25976   2.19741   2.12993   0.06747  -0.37910
;
/* Comentarios adicianais */
```

- Set Dir1: Indica a base de dados para diretoria 1.
- Put Ind1-ind49 : Está indicando que vai juntar a resposta dos 49 indivíduos com a solução do MDS, isto é, a nova base de dados denominado por MODDIR1 vai estar composta pelas respostas dos 49 indivíduos e as dimensões dadas pela solução do MDS.
- dim1: O primeiro atributo encontrado pela solução do MDS.
- dim2: O segundo atributo encontrado pela solução do MDS

- Soma = $dim1^2+dim2^2$.
- dim12 = $dim1^2$.
- dim22 = $dim2^2$.
- dim1dim2=dim1xdim2.

```
/* REGRESSÃO MONÓTONA PARA ESTIMAR O MODELO PONTO IDEAL SIMPLES
        (NÃO MÉTRICO)  DO INDIVÍDUO 1 DA DIRETORIA 1 */

PROC TRANSREG DATA=MODDIR1
        OUTTEST=RESULT1 DAPPROXIMATIONS
          COEFFICIENTSMRC COORDINATES TDPREFIX=COORD
            ADPREFIX=ESTIM;

 MODEL MONOTONE(IND1)=LINEAR(DIM1  DIM2  SOMA);
                        OUTPUT OUT=RESULT2;
  RUN;

PROC PRINT DATA=RESULT1;   RUN;
PROC PRINT DATA=RESULT2;   RUN;


    /* REGRESSÃO MONÓTONA PARA ESTIMAR O MODELO PONTO IDEAL
      PONDERADO (NÃO MÉTRICO) DO INDIVÍDUO 1 DA DIRETORIA 1 */

PROC TRANSREG DATA=MODDIR1 OUTTEST=RESULT3
                DAPPROXIMATIONS COEFFICIENTS
                  MRC COORDINATES TDPREFIX=COORD
                    ADPREFIX=ESTIM;
```

```
 MODEL MONOTONE(IND1)=LINEAR(DIM1  DIM2  DIM12  DIM22);
                        OUTPUT OUT=RESULT4;
   RUN;
PROC PRINT DATA=RESULT3;  RUN;
PROC PRINT DATA=RESULT4;  RUN;


 /* REGRESSÃO MONÓTONA PARA ESTIMAR O MODELO PONTO IDEAL GERAL
          (NÃO MÉTRICO) DO  INDIVÍDUO 1 DA DIRETORIA 1 * /

PROC TRANSREG DATA=MODDIR1 OUTTEST=RESULT5
                 DAPPROXIMATIONS COEFFICIENTS
                     MRC COORDINATES TDPREFIX=COORD
                       ADPREFIX=ESTIM;

 MODEL MONOTONE(IND1)=LINEAR(DIM1  DIM2  DIM12  DIM22  DIM1DIM2);
                        OUTPUT OUT=RESULT6;
   RUN;

PROC PRINT DATA=RESULT5; RUN;
PROC PRINT DATA=RESULT6; RUN;

/* Comentários Adicionais */
```

- Para fazer a estimação da parte métrica, substituir no modelo a palavra MONOTONE por LINEAR.
- Nota-se que o programa foi feito para um indivíduo, também pode-se fazer uma análise considerando a todos os indivíduos de uma mesma diretoria. A única mudança é fazer a substituição de IND1 por (IND1-IND49) no Modelo.

PROGRAMA 8

Este programa foi elaborado com base no "PROC TRANSREG" do SAS. A diferença com o programa 7, é que agora podemos calcular diretamente os Pontos Ideais dos indivíduos. Por exemplo, considerando todos os indivíduos da diretoria 1 temos:

```
/* UMA FORMA DE JUNTAR ARQUIVOS E ADICIONAR VARIÁVEIS */

DATA MODDIR1;
SET DIR1;
PUT IND1-IND49;
INPUT DIM1  DIM2;
CARDS;
 1.73514   0.92212
 0.68235   0.23358
-1.79318   0.46527
 1.52092  -0.60518
-0.25305  -0.98792
-0.43274  -0.28762
-1.45943   0.25976
;
/* USANDO TRANSREG PARA ESTIMAR OS PONTOS IDEAIS SIMPLES PARA
                DIRETORIA 1 (NÃO MÉTRICO) */

PROC TRANSREG DATA=MODDIR1;
        MODEL MONOTONE(COL1-COL49)=POINT(DIM1-DIM2);
                OUTPUT TSTANDARD=CENTER COORDINATES
                        REPLACE NOSCORES OUT=RESULT1;
  RUN;
```

```
/* USANDO TRANSREG PARA ESTIMAR OS PONTOS IDÉIAS PONDERADOS DA
                    DIRETORIA 1 (NÃO MÉTRICO) */

PROC TRANSREG DATA=MODDIR1;
    MODEL MONOTONE(COL1-COL49)=EPOINT(DIM1-DIM2);
                OUTPUT TSTANDARD=CENTER COORDINATES
                    REPLACE NOSCORES OUT=RESULT2;
  RUN;

PROC PRINT DATA=RESULT2;
  RUN;

    /* USANDO TRANSREG PARA ESTIMAR OS PONTOS IDEAIS GERAL PARA
                    DIRETORIA 1 (NÃO MÉTRICO) */

PROC TRANSREG DATA=MODDIR1;
        MODEL MONOTONE(COL1-COL49)=QPOINT(DIM1-DIM2);
                OUTPUT TSTANDARD=CENTER COORDINATES
                    REPLACE NOSCORES OUT=RESULT3;
  RUN;

PROC PRINT DATA=RESULT3;
  RUN;

/* Comentários Adicionais *

- MONOTONE: Indica parte não métrica.
- LINEAR: Indica parte métrica.
- POINT: Indica que vai estimar os Pontos Ideais Simples.
- EPOINT: Indica que vai estimar os Pontos Ideais Ponderados.
- QPOINT: Indica que vai estimar os Pontos Ideais Geral.
```

Por exemplo, a saída do Modelo ponto Ideal Ponderado (considerando só os Pontos Ideais Ponderados) para a diretoria 1 foi a seguinte:

OBS	COL	COL	INTERCEP	DIM1	DIM2	DIM1_2	DIM2_2
1	.	.	.	-0.7640	0.09498	0.33213	4.21466
2	.	.	.	-1.3475	0.22761	0.53451	-1.66663
3	.	.	.	4.5397	-0.17377	-0.17664	-3.20319
4	.	.	.	-0.6933	0.98878	0.91723	0.72698
5	.	.	.	-0.4009	-0.64899	1.48766	-2.53604
6	.	.	.	-0.5252	-0.63726	1.20378	-2.63110
7	.	.	.	-0.5727	0.21971	0.59788	3.30313
8	.	.	.	3.1559	-0.39652	0.08163	-3.12064
9	.	.	.	1.0259	-0.17977	-0.46995	-4.50233
10	.	.	.	1.2820	0.15843	-0.58696	-0.30739
11	.	.	.	-10.2060	-0.81601	0.05807	1.07181
12	.	.	.	-0.0116	-0.81386	-1.75951	0.45552
13	.	.	.	1.5964	-0.21568	-0.15876	-4.94050
14	.	.	.	0.0052	0.81466	1.57455	2.00116
15	.	.	.	-0.3553	-0.61403	1.53425	-2.90694
16	.	.	.	-0.7579	-0.24895	0.78792	0.69700
17	.	.	.	-0.3026	-0.54671	1.15904	-0.63591
18	.	.	.	0.5249	-0.18150	-0.82808	3.39433
19	.	.	.	-1.54377	0.21046	0.28547	2.91753
20	.	.	.	0.72009	0.05579	-0.83392	0.61427
21	.	.	.	0.52058	-0.72445	-0.99073	-0.51533
22	.	.	.	0.20006	-0.09652	-0.66531	5.20954
23	.	.	.	0.88683	1.17104	-0.75876	-0.97640
24	.	.	.	0.15633	-0.29296	-1.22473	3.80220
25	.	.	.	0.06058	-0.09790	-0.56083	5.67188
26	.	.	.	8.13317	-0.82673	-0.08180	0.88622
27	.	.	.	-0.44804	1.49625	1.09945	0.89788
28	.	.	.	0.24928	-2.89314	-1.48471	0.62752
29	.	.	.	0.68553	-0.61064	-0.49255	2.54748
30	.	.	.	3.15589	-0.39652	0.08163	-3.12064
31	.	.	.	-0.60403	-1.01112	1.10660	-0.41705
32	.	.	.	-0.36226	-0.65841	1.47730	-2.79468
33	.	.	.	-9.07483	-0.81936	0.06557	1.05533
34	.	.	.	-3.49320	0.13355	0.22480	-1.74187
35	.	.	.	-2.49972	-2.23802	0.29767	0.09634
36	.	.	.	3.15589	-0.39652	-0.08163	3.12064
37	.	.	.	-3.8745	0.14142	0.20204	-1.46217
38	.	.	.	11.9208	0.21723	-0.06632	-2.22649
39	.	.	.	0.9137	-0.13667	-0.73053	-2.86187
40	.	.	.	1.9063	0.37514	-0.37034	0.43631
41	.	.	.	0.6672	0.03068	-0.85416	-4.58717
42	.	.	.	-0.5693	0.21937	0.60060	3.30541
43	.	.	.	-1.8428	-0.50337	0.19764	2.17276
44	.	.	.	-3.1859	0.71963	0.20440	-1.17249
45	.	.	.	-1.3650	-2.58217	0.51269	0.09381
46	.	.	.	-0.3972	-0.58554	1.45225	-2.98400
47	.	.	.	-1.1820	0.03296	0.50131	1.90156
48	.	.	.	3.1559	-0.39652	-0.08163	3.12064
49	.	.	.	-1.8428	-0.50337	0.19764	2.17276

REFERÊNCIAS BIBLIOGRÁFICAS

ARTES, R. (1991). *Análise de preferência*. São Paulo : IME-USP. 188p. (Dissertação de mestrado em Estatística)

BACHTEL, G. G., TUCKER, L. R. & CHANG, W. A. (1971). A scalar product model for multidimensional scaling of choice. *Psychometrika*, 36, 369-388p.

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. & BRUNK, H. D. (1972). *Statistical inference under order restrictions* : The theory and application of isotonic regression. London : John Wiley. 388p.

BLOXOM, B. (1968). Individual differences in multidimensional scaling. Princeton, N. J. Educational testing Service. *P Research Bulletin*, 45-68p.

BLOXOM, B. (1974). An alternative method of fitting a model of individual differences in multidimensional scaling. *Psychometrika*, 39, 365-367p.

CAMBILLO, M. E. (1996). *Técnicas multivariantes en la investigação comercial*. Lima-Perú : Universidad Mayor de San Marcos-Facultad de Ciencias Matemáticas-Escuela Academico Profesional de Estadística. 1-40p. (cursillo).

CARROLL, J. D. (1972). *Individual differences and multidimensional scaling*. In: Shepard, R. N. Romney, A. K. and Nerlone, S. B., eds. Multidimensional Scaling : Theory and applications in the behavioral sciences. New York : Seminar Press. 105-157p. Vol. 1.

CARROLL, J. D. & CHANG, J. J. (1970). Analysis of Individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" descomposition. *Psychometrika*, 35, 283-319p.

CARROLL, J. D. (1989). A Quase-nonmetric method for multidimensional scaling via an extended euclidean model. *Psychometrika*, 54, 217-229p.

CARROLL, J. D. (1995). Fitting an extended INDSCAL model to Three-way proximity data. *Journal of Classification*, 12, 57-71p.

COOMBS, C. H. (1964). *A theory of data*. New York : John Wiley.

DAVISON, M. L. (1983). *Multidimensional scaling*. New York : John Wiley. 242p.

DE LEEUW, J., YOUNG, F. W. & TAKANE, Y. (1976). Additive struture in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika*, 41, 471-503p.

DILLON, W. R. & GOLDSTEIN, M. (1984). *Multivariate analysis*: Methods and applications. New York : John Wiley & Sons, Inc. 587p.

DRAPER, N. R. & SMITH, H. (1981). *Applied regression analysis*. 2ed. New York : JohnWiley. 709p.

ECKART, C. & YOUNG, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1. 211-218p.

FALAHEE, M. & MACRAE, A. W. (1995). Consumer appraisal of drinking water: multidimensional scaling analysis. *Food Quality and Preference*, 6, 327-332p.

GABRIEL, K. R. & ZAMIR, S. (1979). Lower rank approximation of matrices by least squares with any choice of weights. *Technometrics*, 21, 489-498p.

GOWER, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53, 325-338p.

GREEN, P. E. & TULL, D. S. (1978). *Research for marketing decisions*. 4ed. Englewood Cliffs : Prentice Hall, Inc. 673 p.

GREENACRE, M. J. (1984). *Theory and applications of correspondence analysis*. New York : Academic Press. 364p.

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

GREENACRE, M. J. & UNDERHILL, L. G. (1982). Scaling a data matrix in a low-dimensional euclidean space. In: Hawkins, D. M : *Topics in Applied Multivariate Analysis*. Cambridge : Cambridge University Press. 183-268p.

GROENEN, P. J. F., MATHAR, R. & HEISER, W. J. (1995). The majorization approach to multidimensional scaling for minkowski distances. *Journal of Classification*, 12, 3-19p.

HORAN, J. W. (1969). Multidimensional scaling: Combining observations when individuals haven different perceptual structures. *Psychometrika*, 34. 139-165p.

JAIN, A. K. & DUBES, R. (1988). *Algorithms for clustering data* : Data representation. Englewood Cliffs : Prentice-Hall, Inc. 320p.

JACKSON, J. E. (1991). *A user's guide to principal components*. New York : John Wiley & Sons, Inc. 569p.

JOLLIFE. I. T. (1986). *Principal component analysis*. New York : Springer-Verlag.270p.

JOHNSON, R. A. & WICHERN, D. W. (1982). *Applied multivariate statistical analysis*. Englewood Cliffs : Prentice-Hall. 607p.

KRZANOWSKI, W. J. & MARRIOTT, F. H. C. (1994). *Multivariate analysis part 1* : Distributions, ordination and Inference. New York : John Wiley & Sons, Inc. 280p.

KRZANOWSKI, W. J. & MARRIOTT, F. H. C. (1995). *Multivariate analysis part 2* : Classification, covariance structures and repeated measurements. New York : John Wiley & Sons, Inc. 280p.

KRUSKAL, J. B. (1964a). Multidimensional scaling by optimizing goodness-of-fit to a Non-metric Hypothesis. *Psychometrika*, 29, 1-27p.

KRUSKAL, J. B. (1964b). Nonmetric multidimensional scaling: A Numerical method. *Psychometrika*, 29, 115-129p.

LESTER, A. N., (1969). The use of nonmetric multidimensional scaling in marketing analysis. *Journal of Marketing*, 33, 37-43p.

LEVY, M. M. S. (1981). *Escalonamento multidimensional*. São Paulo : IME-USP. 255p. (Dissertação de mestrado em Estatística).

MANLY, B. F. J., (1994). *Multivariate statistical methods* : A primer. 2º ed. London : Chapman & Hall. 215p.

MARDIA, K. V., KENT, J. T. & BIBBY, J. M. (1979). *Multivariate analysis*. New York : Academic Press. 521p.

RICHARDSON, M. V. (1938). Multidimensional psychophysics. *Psycholigical Bulletin*, 35, 659-660p.

ROUSSEEUW, P. J. & MOLENBERGHS, G. (1993). Transformation of non positive semidefinite correlation matrices. *Commun. Statist*. Theory Meth., 22(4), 965-984p.

SAMMON, J. W. (1969). A non-linear mapping for data structure analysis. *IEEE Transactions in computing*, 401-409p.

SEBER, G. A. F. (1984). *Multivariate observations*. New York : Jhon Wiley & Sons, Inc. 686p.

SCHIFFMAN, S. S., REYNOLDS, M. L. & YOUNG, F. W. (1981). *Introduction to multidimensional scaling*. New York : Academic Press. 413p

SHEPARD, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function I. *Psychometrika*, 27, 125-140p.

SHEPARD, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function II. *Psychometrika*, 27, 219-246p.

SHEPARD, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 39, 373-421p.

SIBSON, R. (1978). Studies in the robustness of multidimensional scaling: Procrustes statistics. *Journal of the Royal Statistical Society*, Series B 40, 234-238p.

TAKANE, Y. (1987). Analysis of contingency tables by ideal point discriminant analysis. *Psychometrika*, 52, 439-513p.

TAKANE, Y., YOUNG, F. W. & DE LEEUW, J. (1976). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimum scaling features. *Psychometrika*, 42, 7-67p.

TAKANE, Y., YOUNG, F. W. & DE LEEUW, J. (1980). An individual differences additive model: An alternating least squares method with optimal scaling features. *Psychometrika*, 45, 183-209p.

TAKEUCHI, K., YANAI, H. & MUKHERJEE, B. N. (1984). *The foundations of multivariate* : A unified approach by means of projection onto linear subspaces. New Delhi : Wiley Eastern Limited. 458p.

TORGERSON, W. S. (1952). Multidimensional scaling I: Theory and method. *Psychometrika*, 17, 401-419p.

TORGERSON, W. S. (1958). *Theory and methods of scaling*. New York : John Wiley and Sons, Inc.

TUCKER, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, 23, 111-136p.

TUCKER, L. R. (1972). Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika*, 37, 3-27p.

WARREN, F. K. (1992). Marketing research: Uncovering competitive advantages. *SAS Institute Inc*. Cary N. C. 1-9p.

YOUNG, F. W, DE LEEUW, J. & TAKANE, Y.(1976). Regression with qualitative and quantitative variables an alternating least squares method with optimal scaling features. *Psychometrika*, 41, 505-529p.

YOUNG, G. & HOUSEHOLDER, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 3, 19-22p.