

EMERSON MARCOS FURTADO



**AUTOMAÇÃO DO RANQUEAMENTO QUALITATIVO
DE ÁREAS ESPECIALMENTE PROTEGIDAS DO
ESTADO DO PARANÁ ATRAVÉS DA
ANÁLISE FATORIAL**

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia do Setor de Ciências Exatas, Área de Programação Matemática da Universidade Federal do Paraná, como requisito parcial à obtenção do Grau e Título de "Mestre em Ciências".

Orientador: Prof. Dr. Anselmo Chaves Neto

CURITIBA
1999

EMERSON MARCOS FURTADO

**AUTOMAÇÃO DO RANQUEAMENTO QUALITATIVO
DE ÁREAS ESPECIALMENTE PROTEGIDAS DO
ESTADO DO PARANÁ ATRAVÉS DA
ANÁLISE FATORIAL**

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia do Setor de Ciências Exatas, Área de Programação Matemática da Universidade Federal do Paraná, como requisito parcial à obtenção do Grau e Título de "Mestre em Ciências".

Orientador: Prof. Dr. Anselmo Chaves Neto

CURITIBA

1999

EMERSON MARCOS FURTADO

**AUTOMAÇÃO DO RANQUEAMENTO QUALITATIVO DE ÁREAS
ESPECIALMENTE PROTEGIDAS DO ESTADO DO PARANÁ
ATRAVÉS DA ANÁLISE FATORIAL**

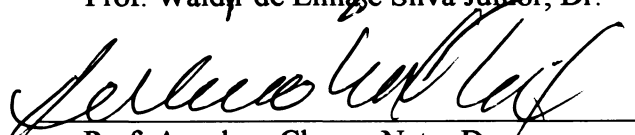
Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciências no Curso de Pós Graduação em Métodos Numéricos em Engenharia da Universidade Federal do Paraná pela comissão formada pelos professores:

Coordenador do Curso:

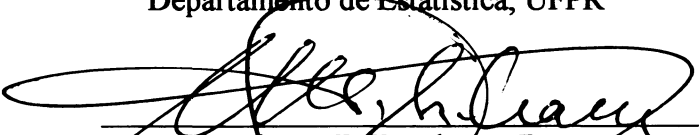


Prof. Waldyr de Lima e Silva Júnior, Dr.

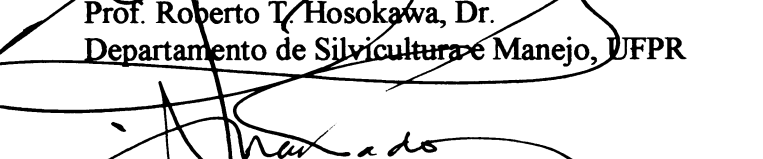
Orientador:



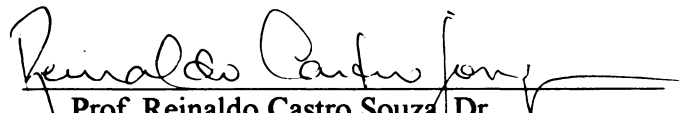
Prof. Anselmo Chaves Neto, Dr.
Departamento de Estatística, UFPR



Prof. Roberto T. Hosokawa, Dr.
Departamento de Silvicultura e Manejo, UFPR



Prof. Roberto Dalledone Machado, Dr.
Departamento de Engenharia Civil, UFPR



Prof. Reinaldo Castro Souza, Dr.
Examinador externo
Departamento de Engenharia Elétrica, PUC-RJ

Curitiba, 12 de novembro de 1999

AGRADECIMENTOS

À cidade de Curitiba, que transformou minha vida, com sua gente, seus lugares e seus mistérios. Aquela que quando aqui cheguei, fez com que eu valorizasse a amizade, o trabalho e a natureza.

À Universidade Federal do Paraná, que proporcionou a oportunidade de estudar mais e desenvolver novas técnicas, através do Curso de Pós-Graduação em Métodos Numéricos em Engenharia.

Ao meu orientador Prof. Dr. Anselmo Chaves Neto. Admirável amigo, que tanto me incentivou e me apoiou desde o início da jornada. Sua experiência, lucidez e orientação segura foram imprescindíveis ao trabalho.

Ao Fábio Adriano Merlin Skrobot, pela sua paciência, amizade e auxílio nos recursos de informática. Nos finais de semana, nas horas mais impróprias, sempre pude contar com o seu apoio.

Ao amigo Prof. Dr. Rubens Robles Ortega, que, além de professor, quando tudo ainda eram intenções, ajudou-me direcionando caminhos e apresentando pessoas que muito me auxiliaram.

À eng.^a Florestal Zilna Hoffmann Domingues, pessoa da qual surgiu a idéia de se realizar este trabalho. Pela sua boa vontade, comparecendo a reuniões, fornecendo material de pesquisa e clareando meus caminhos.

Ao eng.^o Agrônomo Wilson Loureiro, que disponibilizou tempo para me ouvir, forneceu-me material de pesquisa e sempre esteve pronto a me ajudar esclarecendo questões sobre Faxinais.

Aos professores do Curso de Pós Graduação em Métodos Numéricos em Engenharia, da Universidade Federal do Paraná, na Área de Programação Matemática, especialmente à prof.^a Dra. Teresinha, ao prof. Dr. Celso, à prof.^a MSc Beatriz e à prof.^a Dra. Neida, pelo conhecimento transmitido e boa vontade em todos os esclarecimentos.

À Fabiana, namorada e companheira, que tanto me incentivou, pela sua compreensão nas minhas ausências. Sua força muito me ajudou.

Aos meus pais, que me deram oportunidade de estar aqui.

Aos meus fiéis amigos, Bruno, Carlos e Marcelo, que sempre me incentivaram nas dificuldades e me apoiaram em tudo.

A todos que de alguma forma, direta ou indiretamente, me ajudaram neste trabalho.

SUMÁRIO

LISTA DE QUADROS	vii
LISTA DE FIGURAS	viii
LISTA DE ANEXOS	ix
RESUMO	x
ABSTRACT	xi
1 INTRODUÇÃO	1
1.1 APRESENTAÇÃO DO PROBLEMA	1
1.2 OBJETIVOS	2
1.3 ESTRUTURA DO TRABALHO	3
2 REVISÃO DE LITERATURA	7
2.1 ICMS ECOLÓGICO	7
2.2 ÁREAS ESPECIALMENTE PROTEGIDAS	9
2.3 SISTEMA FAXINAL	11
2.3.1 FAXINAL	12
2.3.2 HISTÓRICO	13
2.3.3 PROCESSO DE DESAGREGAÇÃO	16
2.4 ANÁLISE FATORIAL	18
2.4.1 INTRODUÇÃO	18
2.4.2 ANÁLISE FATORIAL x ANÁLISE DE COMPONENTES PRINCIPAIS	26
2.4.3 MODELO FATORIAL ORTOGONAL	33
2.4.4 SINGULARIDADE DOS PARÂMETROS	40
2.4.5 NÚMERO DE FATORES COMUNS DA ANÁLISE FATORIAL	43
2.4.6 MÉTODOS DE ESTIMAÇÃO	58
2.4.6.1 Método de Componentes Principais	59

2.4.6.2 Verossimilhança da Normal Multivariada	63
2.4.6.3 Método da Máxima Verossimilhança	67
2.4.6.4 Método do Fator Principal	70
2.4.6.5 Outros Métodos de Estimação	73
2.4.7 ROTAÇÃO DE FATORES	75
2.4.7.1 Critério Quartimax	81
2.4.7.2 Critério Varimax	85
2.4.7.3 Varimax Normal	87
2.4.7.4 Rotação Varimax Proposta por Kayser	92
2.4.7.5 Rotação Varimax Proposta por Horst	104
2.4.7.6 Método Promax	110
2.4.8 ESCORES FATORIAIS	112
2.4.8.1 Método dos Mínimos Quadrados Ponderados	114
2.4.8.2 Método da Regressão	118
2.4.8.3 Método da Regressão para Fatores Correlacionados	127
2.4.9 PERSPECTIVAS E UMA ESTRATÉGIA PARA A ANÁLISE FATORIAL	130
3 MATERIAIS E MÉTODOS	132
3.1 COLETA E PREPARAÇÃO DOS DADOS	132
3.1.1 AMOSTRAGEM	133
3.1.2 ELABORAÇÃO DO QUESTIONÁRIO	134
3.1.3 DEFINIÇÃO DAS VARIÁVEIS	134
3.2 MÉTODO ESTATÍSTICO DE ANÁLISE	141
3.3 PROGRAMA	143
3.3.1 DESCRIÇÃO	144
3.3.2 FLUXOGRAMA	147

3.3.3	CÁLCULO DOS ESCORES PADRONIZADOS.....	149
3.3.4	MÉTODO HOUSEHOLDER.....	150
3.3.5	MÉTODO QL.....	154
3.3.6	CÓDIGOS.....	158
3.3.7	APLICAÇÃO AOS FAXINAIS.....	158
4	RESULTADOS E CONCLUSÕES.....	164
	ANEXOS.....	166
	REFERÊNCIAS BIBLIOGRÁFICAS.....	219

LISTA DE QUADROS

QUADRO 1 -	CRITÉRIOS E PERCENTUAIS UTILIZADOS PARA RATEIO DO ICMS A QUE OS MUNICÍPIOS TÊM DIREITO NO ESTADO DO PARANÁ.....	7
QUADRO 2 -	CONTRASTES ENTRE ANÁLISE FATORIAL E ANÁLISE DE COMPONENTES PRINCIPAIS	33
QUADRO 3 -	LISTAGEM DOS FAXINAIS ESTUDADOS POR MUNICÍPIO.....	132
QUADRO 4 -	LISTAGEM DO NÚMERO DE FAMÍLIAS, NÚMERO DE AMOSTRAS NECESSÁRIAS E TOMADAS	142
QUADRO 5 -	AUTOVALORES MAIORES QUE 1 DA MATRIZ DE CORRELAÇÃO E PERCENTUAL DA VARIÂNCIA EXPLICADA	159
QUADRO 6 -	COMUNALIDADES E VARIÂNCIAS ESPECÍFICAS DAS VARIÁVEIS.....	160
QUADRO 7 -	MATRIZ DE ESCORES FATORIAIS	162
QUADRO 8 -	RANQUEAMENTO DOS FAXINAIS ANALISADOS.....	163

LISTA DE FIGURAS

FIGURA 1 -	ESQUEMA DA DISPOSIÇÃO FÍSICA DO SISTEMA FAXINAL.....	12
FIGURA 2 -	TESTE <i>SCREE</i>	46
FIGURA 3 -	ROTAÇÃO DE FATORES.....	77
FIGURA 4 -	VARIMAX NORMAL.....	89
FIGURA 5 -	CICLO TRIGONOMÉTRICO.....	97
FIGURA 6 -	INTERFACE DE IDENTIFICAÇÃO.....	145
FIGURA 7 -	INTERFACE DE LOCALIZAÇÃO DA UNIDADE DE CONSERVAÇÃO.....	146
FIGURA 8 -	INTERFACE DE QUALIFICAÇÃO DO FAXINAL.....	146
FIGURA 9 -	INTERFACE DE RESPONSABILIDADE TÉCNICA DO AVALIADOR.....	147
FIGURA 10 -	FLUXOGRAMA DO CÁLCULO DO ESCORE PADRONIZADO.....	148

LISTA DE ANEXOS

ANEXO 01 – QUESTIONÁRIO DO FAXINAL.....	167
ANEXO 02 – QUESTIONÁRIO UNIDADE FAMILIAR.....	172
ANEXO 03 – MATRIZ DE DADOS ORIGINAIS.....	185
ANEXO 04 – MATRIZ DE CORRELAÇÃO.....	191
ANEXO 05 – MATRIZ DE CARREGAMENTOS FATORIAIS.....	200
ANEXO 06 – MATRIZ DE CARREGAMENTOS FATORIAIS ROTACIONADOS.....	203
ANEXO 07 – CÓDIGOS.....	206
ANEXO 08 – FUNÇÃO DE AJUDA DO PROGRAMA.....	217

RESUMO

Os municípios do Paraná que possuem em seus territórios Áreas Especialmente Protegidas, têm, através da Lei Complementar N.º 59/91, conhecida como Lei do ICMS Ecológico, direito a um percentual dos recursos do Imposto sobre Circulação de Mercadorias e Serviços – ICMS. A avaliação de cada área é feita pelo Instituto Ambiental do Paraná – IAP, que realiza anualmente uma avaliação da qualidade, determinando a cada uma delas um escore. Esse escore determina o repasse financeiro ao município. Desta forma, é proposta no presente trabalho uma técnica estatística multivariada, denominada Análise Fatorial, que, através de um programa computacional automático, realize o ranqueamento qualitativo de áreas de Faxinais no Estado. Os Faxinais constituem um tipo especial de área protegida. Os dados brutos usados como entrada no programa, foram obtidos de dois questionários : um deles aplicado as unidades familiares; e outro, as lideranças locais. Foram analisadas 80 variáveis de 20 Faxinais observados, referentes ao ano agrícola de agosto de 1997 a julho de 1998. Utilizando a Análise Fatorial, conseguiu-se a redução das 80 variáveis para apenas 17 fatores. Os fatores considerados contribuíram para uma explicação de 97,7649% do total da variância amostral. O programa elaborado permite ao usuário o ranqueamento automático das áreas, com opções para armazenar escores de anos anteriores e observar a evolução da qualidade de vida de cada comunidade. Pode-se assim, com maior eficiência, balizar as ações de impacto, procurando sempre melhorar a cada ano.

ABSTRACT

The municipal districts of Paraná that possess in its territories Especially Protected Areas, they have, through the Complemental Law N.º 59/91, known as Law of Ecological ICMS, the right to a percentile of the resources of the Tax on Circulation of Goods and Services - ICMS. The evaluation of each area is made by the Environmental Institute of Paraná - IAP, that accomplishes an evaluation of the quality annually, determining each one of them a score. This score determines the reviews financial to the municipal district. Therefore, it is proposed in the present work a multivariate statistical technique, denominated Factor Analysis, that, through a computacional automatic program, accomplish the qualitative ranking of areas of Faxinais in the state. Faxinais constitutes a special type of protected area. The gross data used as entrance in the program, they have been obtained of two questionnaires: one of them applied to family units; and the other, the local leaderships. They have been analyzed 80 variables of 20 observed Faxinais, refering to agriculture year from august 1997 through jully 1998. Using the Factor Analysis, the reduction of the 80 variables was gotten for just 17 factors. The considered factors contributed to an explanation of 97,7649% of the total of the sample variance. The elaborated program allows to the user the automatic ranking of the areas, with options to store scores the previous years and to observe the evolution of the quality in each community's life. Acting this way, it may becon with more efficiency the impact actions, always trying to improve each year.

1 INTRODUÇÃO

1.1 APRESENTAÇÃO DO PROBLEMA

A partir do Decreto Estadual n.º 3.446/97, foram criadas as ARESUR – Áreas Especiais de Uso Regulamentado, abrangendo porções do território do Estado do Paraná caracterizados pela existência dos Faxinais. Os Faxinais foram, assim, a partir de 1997 considerados *Áreas Especialmente Protegidas*.

Na Federação brasileira, cada Estado pode definir um perfil de distribuição dos recursos do ICMS – Imposto sobre Circulação de Mercadorias e Serviços – a que os municípios têm direito. O Estado do Paraná, pioneiramente, definiu pela Lei Complementar n.º 59/91, conhecida como Lei do ICMS Ecológico, que 5% dos recursos do ICMS sejam distribuídos segundo um critério ambiental. Por essa lei, metade dos 5% são distribuídos aos municípios que possuem em seus territórios mananciais, e a outra metade, aos municípios que possuem *Áreas Especialmente Protegidas*, caracterizadas pelas Unidades de Conservação.

Desta forma, os municípios do Estado do Paraná que possuem em seus territórios Faxinais passaram a ter direito de parte dos recursos do ICMS. Entretanto, é necessário que cada município mantenha um satisfatório nível de qualidade de vida nesses Faxinais.

O Instituto Ambiental do Paraná – IAP realiza anualmente a avaliação da qualidade de vida nas *Áreas Especialmente Protegidas*. A avaliação permite que seja atribuída à área um score ou nota, que determinará ou não um incremento ou repasse financeiro ao município. Desta forma, o município pode ser beneficiado financeiramente não só pela existência de *Áreas Especialmente Protegidas*, mas também pelo nível de qualidade da área protegida.

Pelo decreto, as variáveis analisadas são a densidade populacional, a qualidade de vida das populações residentes, a organização comunitária, o nível de comprometimento e o empenho dos municípios para com o desenvolvimento social e econômico das áreas.

Assim sendo, a proposta do presente estudo é fazer um ranqueamento de *Áreas Especialmente Protegidas*, considerando as variáveis avaliadas pelo IAP. Visando o ICMS Ecológico, as áreas consideradas foram os Faxinais inscritos no Cadastro Estadual de Unidades de Conservação.

Como solução para se medir todas as variáveis propostas e realmente ranquear os Faxinais, é utilizada a Análise fatorial, técnica da Análise Multivariada. Desta forma, é construído um programa computacional automático, que a partir dos dados brutos compute as variáveis e determine um escore para cada Faxinal avaliado.

1.2 OBJETIVOS

Tendo em vista o decreto n.º 3.446/97, que reconheceu os Faxinais e instituiu as ARESUR(s), dando oportunidade de remuneração através da Lei do ICMS Ecológico, este estudo objetiva:

- a) criar critérios para um ranqueamento de Unidades de Conservação, visando ou não a adesão ao ICMS Ecológico;
- b) o desenvolvimento de um programa que simplifique as entradas de dados de qualquer tipo de *Área Especialmente Protegida*;
- c) o desenvolvimento de um programa que atualize os dados das ARESUR(s), atendendo às necessidades dos serviços públicos responsáveis;
- d) fazer uma revisão detalhada da Análise Fatorial, técnica da Análise Multivariada.

1.3 ESTRUTURA DO TRABALHO

O presente trabalho se encontra estruturado em quatro capítulos: Introdução, Revisão de Literatura, Materiais e Métodos, e Resultados e Conclusões.

No capítulo da Introdução, o problema de ranqueamento de *Áreas Especialmente Protegidas* é apresentado. No caso, as áreas ranqueadas são Faxinais do Estado do Paraná. As variáveis analisadas em cada Faxinal são citadas, e para a solução do problema é proposta a Análise Fatorial. Para atualizar-se ano a ano a classificação qualitativa dos Faxinais é proposto um programa computacional automático. Além disso, neste mesmo capítulo, são expostos os quatro objetivos deste trabalho.

O segundo capítulo, que trata da Revisão de Literatura tem destacados a lei do ICMS Ecológico, as *Áreas Especialmente Protegidas*, o Sistema Faxinal e a Análise Fatorial.

A Lei do ICMS Ecológico beneficia, através da Lei Complementar n.º 63/90, municípios que possuam em seus territórios Unidades de Conservação. Aqui, destacam-se os tipos de áreas que podem ser beneficiadas com a lei e a importância para cada município de sempre procurar aumentar o nível da qualidade de vida das populações que ali residem, conciliando isso com a preservação da natureza. O repasse de recursos financeiros pelo Governo do Estado, dependendo do escore avaliado pelo IAP, representa o prêmio pago para cada município.

Os Faxinais do Paraná são áreas de preservação das coberturas vegetais naturais e uma importante forma de organização camponesa, que vive em comunidade. Desta forma, é realçada ainda no segundo capítulo toda a história dos Faxinais e o processo atual de desagregação do Sistema Faxinal, suas causas e suas conseqüências, assim como a importância ecológica, social e cultural da continuidade desta forma de organização.

A técnica da Análise Multivariada, utilizada no ranqueamento, é apresentada através da Análise Fatorial. Na seção intitulada Análise Fatorial são descritos os seguintes itens: Análise Fatorial \times Análise de Componentes Principais, Modelo Fatorial Ortogonal, Singularidade de Parâmetros, Número de Fatores Comuns, Métodos de Estimação, Rotação de Fatores, Escores Fatoriais e Perspectivas e uma Estratégia para a Análise Fatorial.

Na Introdução são mostrados os campos de aplicações mais recentes da Análise Multivariada e um pouco da história da Análise Fatorial. Na comparação entre a Análise Fatorial e a Análise de Componentes Principais, mostram-se as diferenças e sugerem-se casos em que normalmente se faz opção por uma delas.

É importante ressaltar que, em toda a Análise Fatorial, os resultados descritos são provados logo após os correspondentes assuntos. Assim, se durante o texto, é feita referência a um determinado resultado, a descrição e prova do resultado são encontradas logo após o texto.

O Modelo Fatorial Ortogonal apresenta as variáveis e parâmetros definidos. Dependendo das condições da análise, pode existir uma infinidade de possibilidades para os parâmetros do modelo. Desta forma, na Singularidade de Parâmetros, comenta-se as situações que podem ocorrer no cálculo dos carregamentos fatoriais.

No item Número de Fatores Comuns, são definidos cinco critérios para se estabelecer um número m de fatores satisfatório para a análise que se deseja efetuar. No ranqueamento qualitativo de Áreas Especialmente Protegidas, foi escolhido o critério 2, ou seja, o número de fatores comuns na análise é igual ao número de autovalores maiores que 1 (um) da matriz de correlação.

Ainda na Análise Fatorial, os Métodos de Estimação descritos são o Método de Componentes Principais, o Método da Máxima Verossimilhança e o Método do Fator

Principal. Além disso, comenta-se sobre outros métodos não muito comuns, mas que também podem ser aplicados na estimação da matriz de carregamentos fatoriais. Apesar de o Método de Máxima verossimilhança apresentar resultados extremamente consistentes, neste trabalho o método escolhido de estimação foi o de Componentes Principais.

A Rotação de Fatores descreve o histórico dos principais métodos de rotação até o aparecimento da Rotação Varimax Normal. São descritos o Método Quartimax, o Método Varimax, o já citado Método Varimax Normal, as Rotações Varimax Propostas por Kayser e por Horst e o Método Promax para fatores correlacionados. Na aplicação aos Faxinais, o método de rotação utilizado foi o Método Varimax, proposto por Kayser.

Os valores estimados dos fatores comuns são os escores fatoriais. Os métodos descritos para a obtenção dos escores fatoriais são o Método dos Mínimos Quadrados Ponderados, o Método da Regressão e o Método da Regressão para Fatores Correlacionados. No presente trabalho utilizou-se o Método dos Mínimos Quadrados Ponderados.

Nas Perspectivas e uma Estratégia para a Análise Fatorial é sugerido um procedimento estratégico para uma melhor análise da estrutura de dados.

O capítulo 3 trata de Materiais e Métodos, e está dividido em Coleta e Preparação dos Dados, Método Estatístico de Análise e Programa.

A Coleta e Preparação dos Dados lista os 20 Faxinais efetivamente analisados da pesquisa e divide-se em Amostragem, Elaboração do Questionário e Definição das Variáveis. Os dados coletados referem-se ao ano agrícola de agosto de 1997 a julho de 1998. Foram aplicados dois questionários, sendo um à unidade familiar; e outro, às lideranças locais. As variáveis foram sugeridas por DOMINGUES, considerando essencialmente o Decreto Estadual n.º 3.446/97.

O Programa foi desenvolvido em linguagem de programação Visual Basic. O caminho percorrido desde os dados brutos até o escore final padronizado está esquematicamente organizado em um fluxograma. A partir da entrada dos dados são calculadas as correlações entre cada par de variáveis. Em seguida, são obtidos os autovalores e autovetores da matriz de correlação que produzem a matriz de carregamentos fatoriais. Com o objetivo de melhorar a interpretação dos fatores, essa matriz então é rotacionada pelo Método Varimax proposto por Horst. Os escores fatoriais são obtidos pelo Método dos Mínimos Quadrados Ponderados. Uma vez que os fatores têm um grau de importância diferente, o escore final médio é ponderado pela importância de cada fator, dependendo do correspondente autovalor. O Programa permite ao usuário definir o intervalo que deseja situar os escores finais. Assim, os escores padronizados são obtidos em função do intervalo definido pelo usuário. No caso dos Faxinais, o IAP determina que os escores finais pertençam ao intervalo que vai de 0 até 2.

Destacam-se também no Programa os Métodos Householder e QL. O primeiro transforma a matriz simétrica de correlação em uma matriz tridiagonal. O segundo encontra os autovalores e autovetores dessa matriz tridiagonal. Desta forma, com os dois métodos atuando sucessivamente, consegue-se um tempo muito menor na computação dos resultados.

Ainda no capítulo 3, no item aplicação aos Faxinais, foram destacados os autovalores da matriz de correlação, o grau de explicação de acordo com o número de fatores, as comunalidades e variâncias específicas, os fatores considerados, a matriz de escores fatoriais, os escores brutos, os escores padronizados e o ranqueamento dos 20 Faxinais observados.

O capítulo 4 refere-se aos Resultados e Conclusões do trabalho.

Os anexos trazem os questionários do Faxinal (Anexo 01) e da unidade familiar (Anexo 02), a matriz de dados originais (Anexo 03), a matriz de correlação entre as variáveis (Anexo 04), a matriz de carregamentos fatoriais (Anexo 05), a matriz de carregamentos fatoriais rotacionados (Anexo 06), os códigos principais da programação (Anexo 07) e a função de ajuda do programa (Anexo 08).

2 REVISÃO DE LITERATURA

2.1 ICMS ECOLÓGICO

A Constituição brasileira estabelece que estados e municípios têm direito aos recursos oriundos do Imposto sobre Circulação de Mercadorias e Serviços – ICMS – rateados segundo critérios estabelecidos por leis. Segundo a lei complementar n.º 62/90, os estados podem definir critérios para a distribuição de apenas 25% dos recursos a que os municípios têm direito. Os 75% restantes devem ser repassados em função do Valor Adicionado que foi definido na Lei Complementar federal n.º 63/90.

O primeiro estado a incluir o critério ambiental para a distribuição do ICMS foi o Paraná. Segundo a Lei Complementar n.º 59/91, conhecida como Lei do ICMS Ecológico, ficou definido que 5% seria repassado ao município pelo critério ambiental.

O Quadro 1 ilustra os critérios de distribuição antes e depois da lei do ICMS Ecológico.

QUADRO 1 – CRITÉRIOS E PERCENTUAIS UTILIZADOS PARA RATEIO DO ICMS A QUE OS MUNICÍPIOS TÊM DIREITO NO ESTADO DO PARANÁ.

CRITÉRIOS	ATÉ 1991	APÓS 1992
Valor adicionado	80%	75%
Valor da produção agropecuária	8%	8%
Número de habitantes do município	6%	6%
Número de propriedades rurais	2%	2%
Superfície territorial do município	2%	2%
Índice igualitário ou fixo	2%	2%
Ambiental (ICMS Ecológico)	----	5%
TOTAL	100%	100%

FONTE: IAP, 1997

De acordo com a Lei, metade dos 5% seriam distribuídos aos municípios do Estado que possuem em seus territórios mananciais de abastecimento público para municípios vizinhos, e a outra metade, aos municípios que possuem Unidades de Conservação.

O índice utilizado atualmente para o cálculo é o Coeficiente de Conservação da Biodiversidade – CCB. LOUREIRO (1998) definiu esse coeficiente como sendo “a relação entre a superfície da Unidade de Conservação (ou espaço especialmente protegido), com qualidade física satisfatória (ou porção em recuperação) e a superfície territorial do município, corrigido pelo fator de conservação, representado por diferentes categorias de manejo de Unidades de Conservação ou modalidades de outras *Áreas Especialmente Protegidas*”.

Pelo Decreto Estadual n.º 3.446/97, foram criadas as ARESUR – Áreas Especiais de Uso Regulamentado, abrangendo porções do território do Estado, caracterizados pela existência dos Faxinais. O Decreto diz ainda que tanto a criação quanto o benefício financeiro, previsto pelo ICMS Ecológico, definido no próprio Decreto, poderá ser feito a partir da manifestação de interesse do município que possuir Faxinais, devendo, para tanto, além da solicitação, apresentar proposta, negociada com as comunidades, das ações a serem desenvolvidas, a partir, dentre outras, das variáveis a serem avaliadas anualmente. As variáveis analisadas são:

- a) densidade populacional;
- b) qualidade de vida das populações residentes, organização comunitária ;
- c) nível de comprometimento e empenho dos municípios para com o desenvolvimento social e econômico dos Faxinais.

2.2 ÁREAS ESPECIALMENTE PROTEGIDAS

Segundo a Lei do ICMS Ecológico, os municípios que possuem em seus territórios *Áreas Especialmente Protegidas* podem obter benefícios no rateio dos recursos financeiros do Imposto sobre Circulação de Mercadorias e Serviços – ICMS.

Entende-se por *Áreas Especialmente Protegidas* as Unidades de Conservação, Áreas de Terras Indígenas, Reservas Particulares do Patrimônio Natural e Reservas Florestais Legais, Áreas de Preservação Permanentes, Sítios Especiais e Faxinais, desde que estejam nos contornos das Unidades de Conservação.

As Unidades de Conservação são as porções do território nacional, incluindo águas territoriais, com características naturais de relevante valor, de domínio público ou privado, legalmente instituídas pelo Poder Público, com objetivos e limites definidos, sob regime especial de administração, as quais se aplicam garantias adequadas de conservação. São Unidades de Conservação as Reservas Biológicas, as Estações Biológicas, os Parques, as Florestas, as Reservas Particulares de Patrimônio Natural e as Áreas de Relevante Interesse Ecológico.

Desta forma, para uma melhor compreensão, pode-se esquematicamente classificar quanto à modalidade, as *Áreas Especialmente Protegidas* em:

- a) Unidades de Conservação;
- b) Áreas de Terras Indígenas;
- c) Sítios Especiais;
- d) Faxinais;
- e) Reservas de Florestas Legais;
- f) Áreas de Preservação Permanente.

Da mesma forma, quanto às categorias de manejo, pode-se classificar as Unidades de Conservação em:

- a) Reservas Biológicas;
- b) Estações Biológicas;
- c) Parques;
- d) Florestas;
- e) Reservas Particulares de Patrimônio Natural;
- f) Áreas de Relevante Interesse Ecológico.

Pelo Decreto Estadual N.º 3.446/97, os Faxinais são entendidos como um sistema tradicional, característico da região Centro-Sul do Paraná, que tem como traço marcante o uso coletivo da terra para a produção animal e a conservação ambiental. Estão fundamentados na integração de três componentes:

- a) produção animal coletiva, à solta, por meio dos criadores comunitários;
- b) produção Agrícola – policultura alimentar de subsistência para consumo e comercialização;
- c) extrativismo florestal de baixo impacto – manejo de erva-mate, araucária e outras espécies nativas.

Esse modo de produção auto-sustentável foi denominado, segundo o decreto, de “Sistema Faxinal”.

2.3 SISTEMA FAXINAL

O Sistema Faxinal constitui-se na forma histórica de organização social da produção que mais preservou as condições ambientais, comparada a outras formas organizacionais da produção. Segundo CHANG (1985), é hoje uma organização camponesa, onde coexistem o uso comum e a posse privada da terra.

Esse sistema economicamente se assenta sobre o seguinte tripé:

- a) criação extensiva de animais em áreas comuns;
- b) extração de erva-mate e secundariamente de madeira, também dentro do criadouro comum;
- c) policultura alimentar nas terras de planta circunvizinhas.

O Sistema Faxinal é dividido em dois grandes grupos de terras : as terras de criação e as terras de planta.

As terras de criação são conjuntos de propriedades privadas e contíguas colocadas em uso comum. São formadas por vales ou áreas mais deprimidas, com relevo suave ondulado e presença de aguadas. Predominam nessas áreas solos vermelhos, ácidos e profundos, favoráveis ao desenvolvimento de espécies vegetais de grande porte.

As terras de planta são propriedades privadas e, embora também contíguas, seu usufruto é privado. Localizam-se em áreas mais íngremes do relevo. Predominam nelas solos brancos e rasos e menos ácidos, favorecendo o desenvolvimento da agricultura. Nesse sistema os agricultores não moram junto às plantações, mas, sim, na área verde organizada coletivamente para a criação de animais. A Figura 1 ilustra a estrutura física do Sistema Faxinal.



FIGURA 1 - ESQUEMA DA DISPOSIÇÃO FÍSICA DO SISTEMA FAXINAL

2.3.1 FAXINAL

Segundo DOMINGUES (1999), Faxinal é um sistema agrossilvipastoril secular com características singulares de uso da terra, sendo uma experiência auto-sustentada de relevante importância ecológica, social, histórica e cultural da região Centro-Sul do Paraná. Foi de

grande importância na primeira metade deste século, principalmente quando aproximadamente 1/5 do território do estado era por ele marcado como forma predominante de organização da produção rural.

A referência da população local vinda dos campos é de que faxinal era um mato grosso, ou seja, um mato mais denso, quando comparado aos campos e matos ralos. Entretanto, do ponto de vista técnico de classificação de matas, faxinal significa mato ralo, com vegetação variegada ou braço de campo com árvores esguias, que penetra as matas. Assim, mesmo denominado de forma incorreta, faxinal representa um mato grosso.

É nesse tipo de mata que foram formados os criadouros comuns. Os colonos então empregaram a palavra faxinal tanto à mata quanto ao criadouro comum, e há fortes evidências de que a palavra faxinal tenha se originado dos caboclos nativos.

2.3.2 HISTÓRICO

Historicamente, há uma sucessão de ciclos econômicos no Paraná. Primeiro a mineração, principalmente com o ouro; depois a pecuária. Ambos tiveram grande importância na fixação do homem e no desenvolvimento inicial da província. Em seguida, vieram o ciclo madeireiro e a erva-mate.

A descoberta do ouro nos riachos da Bacia de Paranaguá, no fim do século XVI, estimulou a primeira ocupação agrícola dos solos da região do litoral e principalmente do Primeiro Planalto curitibano. Foi no Paraná que se iniciou a mineração brasileira, com Gabriel de Lara, que, em 27 de novembro de 1646, declarou, na Câmara de São Paulo, ter encontrado ouro em Paranaguá.

Assim, pouco a pouco foi se desenvolvendo uma economia de subsistência, organizada ao lado da mineração. Nesta época os homens distribuíam-se basicamente em mineradores, agricultores e criadores. Com a redução contínua da quantidade de minério extraído, a mineração assim, torna-se uma atividade antieconômica diante dos custos elevados da busca, extração e fundição. Desta forma, o ciclo da mineração perde força em menos de um século.

A condensação da população na zona de Curitiba, a progressiva decadência da mineração, as ligações mais fortes com o mercado de São Paulo e Minas Gerais e, principalmente, a atração exercida pelo mercado consumidor de carne e gêneros agrícolas no litoral foram fatores que contribuíram para o fortalecimento da pecuária nos Campos Gerais.

Ponta Grossa, Jaguariaíva, Rio Negro, Lapa e outras cidades no Segundo Planalto surgiram de pousadas de tropeiros, que percorriam o Caminho de Viamão, transportando gado do Rio Grande do Sul para a feira de Sorocaba. Nessa época, por se localizar nas proximidades dos caminhos dos tropeiros, a vila de Curitiba passa a ser considerada a nova sede da comarca.

A população que vivia de subsistência se transforma em criadores e agricultores para o mercado interno. A mão-de-obra escrava, que antes era ocupada na mineração, agora passava às fazendas. As condições físico-naturais dos Campos Gerais contribuíram decisivamente para a expansão da atividade pecuária nesta fase.

O fato de haver uma disponibilidade considerável de campos ainda não ocupados facultou às pessoas assalariadas da época a possibilidade de se desligarem das fazendas onde trabalhavam e instalarem-se como agricultores ou criadores autônomos independentes. As pessoas que pertenciam a uma camada intermediária, que não eram nem proprietários nem escravos, eram os denominados agregados. Os agregados eram, na época, pessoas

juridicamente livres, mas inteiramente subordinados à classe senhorial. Residiam com suas famílias nas terras das fazendas em pontos distantes da sede. A denominação agregados é pelo fato de serem pessoas que moravam nas fazendas mesmo sem serem proprietárias.

Os campos foram sendo ocupados e fazendas foram sendo formadas. Mesmo assim, a criação encontrava-se em grande atraso técnico. As propriedades não geravam grande renda. Os membros novos da grande família fazendeira viram-se forçados a procurar outras fontes de renda nas matas. Inicialmente foram em busca das matas mistas da própria região, mais tarde, as matas do sertão e do oeste, e por último, as matas do norte do estado. Eram na maioria pessoas consideradas desclassificadas social e economicamente pela divisão do latifúndio.

A extração da madeira se deu desde o início da colonização. Porém, como fator de ocupação, torna-se importante a partir do século XIX, com a utilização do pinho com a finalidade inicial para construção civil e móveis, e, posteriormente, para as indústrias de papel. Grande parte do vale do Iguaçu foi ocupada pelas serrarias, que também geraram núcleos de ocupação, empregando inclusive mão-de-obra de imigrantes eslavos.

Após a abolição da escravatura, ocorrida em 13 de maio de 1888, imigrantes europeus, principalmente poloneses, instalaram-se na região das Matas Mistas dedicando-se a pecuária e sobretudo à atividade ervateira. Na região das Matas Mistas é que encontravam os mais densos ervais nativos. A economia da erva-mate cresce com isso, a ponto de se tornar o principal produto regional e até estadual. A rentabilidade era superior a qualquer outra forma de cultura de plantio. Paralelamente a atividade ervateira; o tipo de vegetação favorecia a criação de animais à solta, principalmente de suínos. Pode-se afirmar que a grande causa que realmente contribuiu para a ocupação dos Campos Gerais foi a erva-mate. Nesta fase é que começa a se formar o que se denomina Sistema Faxinal.

A importância da erva-mate levou a decisão de que os ervais seriam cercados e preservados. As demais terras poderiam ser ocupadas para a lavoura. Devido ao perigo da criação miúda andar à solta e adentrar nas lavouras, as criações eram limitadas territorialmente por acidentes geográficos, tais como rios ou vales profundos, e as lavouras eram cercadas. Desta forma, a racionalidade da produção levou à diferenciação dos Faxinais em áreas de lavouras e áreas de criação.

As cercas dos criadouros passaram a desempenhar um papel importante na estrutura do Faxinal. A construção das cercas era proporcional à área que cada um possuía no Faxinal, e cada um ficava responsável por uma parte da cerca. As prefeituras começaram a participar ativamente dos criadouros comuns, chegando até a inspecionar, julgar casos e intermediar questões.

Enquanto a erva-mate se mostrava um produto de grande rentabilidade, a comunidade crescia e prosperava.

2.3.3 PROCESSO DE DESAGREGAÇÃO

A queda vertiginosa das exportações da erva-mate em 1930 marcou a grande crise da economia ervateira no Estado. Consequentemente, a principal atividade que até então dava manutenção econômica ao Sistema Faxinal passou a perder espaço. Mesmo assim, contrário às expectativas, o Sistema Faxinal permaneceu ainda de forma bastante sólida nas próximas três décadas. O que mudou foram as atividades de subsistência, que passaram a desempenhar um papel complementar.

O declínio da atividade ervateira, tem como consequência, por herança, um fracionamento das fazendas e das unidades menores de produção. Desta forma, o predomínio numérico das pequenas unidades de produção levou a considerar o Sistema Faxinal como uma forma de organização camponesa.

A partir da década de 60, na tentativa de fortalecimento das indústrias, a política agrícola foi orientada para a utilização mais intensiva de máquinas e fertilizantes, explicitando os interesses industriais e, ao mesmo tempo, valorizando lavouras de exportação, tais como café, soja e milho.

O financiamento da agricultura se torna o principal mecanismo de mudança tecnológica, distribuindo subsídios e subordinando a agricultura à esfera financeira através do crédito rural. Desta forma, para os adeptos das técnicas modernas, o criadouro comum deixa de desempenhar o mesmo papel na reprodução da unidade familiar. Pelo contrário, a área em comum passa a ser alvo de investimentos mais intensivos, em que o pressuposto da apropriação privada dos benefícios do próprio investimento faz com que o proprietário reivindique a privatização do uso da terra.

Além dessa modernização agrícola, outros fatores contribuíram para a desagregação dos Faxinais, tais como o esgotamento da fronteira agrícola, os reflorestamentos, o esgotamento dos recursos naturais, o declínio da produção da erva-mate e a lei dos 4 fios.

O processo de desagregação se dá de forma gradual, através do abandono da criação à solta, iniciando-se pela miúda e estendendo-se gradualmente à graúda, e através da diminuição da área total dos criadouros, devido à saída de alguns proprietários desinteressados em criar à solta.

Atualmente, praticamente todos os Faxinais do Estado do Paraná estão se desagregando. Uns em estágios mais avançados; outros nem tanto. Do ponto de vista

ecológico, o Sistema Faxinal serve como forma de preservação das coberturas vegetais naturais. Já do ponto de vista social, os criadouros comuns permitem que a população dos agregados, os mini e pequenos produtores se viabilizem, enquanto pequena produção, com um mínimo de aplicação de capital e mão-de-obra. Entretanto, a permanência dos Sistemas Faxinais requer que se tornem viáveis do ponto de vista econômico.

2.4 ANÁLISE FATORIAL

2.4.1 INTRODUÇÃO

Até a metade do século XX, a aplicação de qualquer método iterativo demandava um enorme esforço de tempo e matemáticos para ser completado. Muito da Ciência conhecida não era utilizado na prática devido às dificuldades na computação dos resultados. O esforço anglo-americano na busca de desenvolvimento no período entre a Primeira e Segunda Grandes Guerras Mundiais, propiciaram a construção de uma máquina que veio alterar muito da vida humana. Essa máquina é hoje conhecida como computador eletrônico.

Do primeiro computador a válvulas até os mais modernos computadores atuais, muito progresso científico foi feito. Com o avanço tecnológico, as técnicas que na época eram consideradas inviáveis pela grande quantidade de cálculos passaram a ser mais utilizadas. A velocidade de processamento foi crescendo, e com ela, os custos operacionais caíam para valores bem menores. Algumas técnicas estatísticas baseadas em métodos recursivos aguardaram muitas décadas até que pudessem se desenvolver plenamente. Dentre estas

técnicas, uma que encontrava-se completamente construída teoricamente por volta da década de 30, mas só teve aplicação plena no final dos anos 50 com a computação eletrônica, foi a Análise Fatorial .

A Análise Fatorial é uma técnica estatística multivariada, que trata do relacionamento entre conjuntos de variáveis. Inicialmente foi desenvolvida por psicólogos, como Spearman, Thomson, Thurstone e Burt, como uma técnica pioneira, tratando de hipóteses sobre a organização da habilidade mental, sugerida pelo exame das matrizes de correlação ou covariância para conjuntos de testes de variáveis cognitivas. O trabalho precoce neste difícil campo gerou permanentes controvérsias pelo lado da psicologia, que por muito tempo desencorajou o interesse mostrado por matemáticos nos problemas teóricos envolvidos. Durante muitos anos o assunto foi a ovelha negra da Teoria Estatística. Gradualmente um mais coerente tratamento do assunto foi desenvolvido, as muitas concepções foram sendo entendidas integralmente e com grande respeitabilidade. Hoje, a Análise Fatorial é uma técnica estatística amplamente usada.

A essencial proposta da Análise Fatorial é descrever, se possível, a estrutura de covariância do relacionamento entre várias variáveis em termos de um número menor de variáveis não observáveis, denominadas fatores. Basicamente, o Modelo Fatorial é motivado pelo fato de que, dentro de um grupo, todas as variáveis têm, entre elas, uma forte correlação. Entretanto, possuem relativamente uma pequena correlação com as variáveis de outro grupo. Assim, cada grupo de variáveis pode representar uma simples estrutura ou fator. O fator fica responsável pelas correlações observadas no grupo. As variáveis seriam, portanto, agrupadas em diversos fatores.

No presente estudo, a Análise Fatorial será aplicada no ranqueamento de Áreas Especialmente Protegidas do Estado do Paraná. A partir de um conjunto inicial de variáveis é

avaliado o grau de relacionamento entre todas as variáveis. A partir daí, com os autovalores e autovetores da matriz de correlação, determina-se o número de fatores considerando-se a porção da variância explicada. Para a estimação dos fatores utilizou-se o Método de Componentes Principais, e para uma melhor interpretação dos resultados utilizou-se a Rotação Varimax Normal. Os escores fatoriais foram estimados pelo Método de Mínimos Quadrados Ponderados.

É de grande conhecimento que, hoje, os métodos multivariados são aplicados em diversos campos de pesquisa. JOHNSON & WICHERN (1998) descrevem com exemplos as principais e mais recentes aplicações da Análise Multivariada. São elas:

a) Medicina e Saúde

Um estudo foi realizado para investigar a reação dos pacientes de câncer sob ação da radioterapia. As medidas foram feitas em 98 pacientes, e direcionadas para 6 tipos principais de observação:

1. número de sintomas
2. quantidade de atividade - escala de 1 a 5
3. tempo de sono - escala de 1 a 5
4. quantidade de alimento consumido - escala de 1 a 3
5. apetite - escala de 1 a 5
6. reação da pele - escala de 0 a 3

A Análise Multivariada foi usada para construir e interpretar as reações dos pacientes à radioterapia, em função de simples parâmetros utilizados em cada tipo de observação.

Respostas específicas referentes a estimulações visuais podem ser registradas através de impulsos elétricos no couro cabeludo humano e analisados por computadores. Em estudos

médicos feitos a partir de efeitos da esclerose múltipla em sistemas visuais, a Análise Multivariada foi usada para examinar os dados da análise visual eletroencefalográfica, classificando e separando os mesmos de forma a identificar as pessoas que sofrem de esclerose múltipla, causada por patologia visual daquelas que não sofrem dessa moléstia.

Programas de planejamento familiar foram implementados em muitos países em desenvolvimento. Um estudo comparou a eficácia de um programa do governo da Malásia, composto de planejamento familiar e saúde, com outros programas exclusivamente familiares, identificando os fatores de integração, que pudessem resultar em uma melhor performance geral. Os resultados do estudo sugeriu uma expansão do programa do governo, integrando os serviços de planejamento familiar com a rede de saúde em vez da expansão de novas organizações exclusivamente de planejamento familiar. A Análise Multivariada identificou os tipos de fatores e formas de integrações das organizações, medindo-os e apontando resultados que pudessem melhorar a performance dessas organizações.

Pesquisas genéticas sobre alcoolismo mostraram que as atividades de duas enzimas são reduzidas significativamente em alcoólatras. O resultado desse estudo foi identificar e medir as variáveis fisiológicas que possam ser usadas efetivamente para discriminar alcoólatras de não alcoólatras.

b) Sociologia

Uma teoria sociológica sugere que a estrutura das ocupações americanas seja determinada por uma forte dimensão sócio econômica. Medidas com 25 variáveis para 583 ocupações foram analisadas usando métodos multivariados para determinar o motivo real dessas ocupações.

A contagem do número de filhos de imigrantes de primeira e segunda geração, nascidos e residentes nos Estados Unidos em 1970 foram tabulados por países de origem e Estado da residência. Métodos multivariados foram usados para verificar a distribuição das nacionalidades dos imigrantes de acordo com o Estado residente.

c) Negócios e Economia

Medidas a partir de diversas variáveis envolvendo finanças e contabilidade, foram realizadas para possibilitar o desenvolvimento de modelos multivariados em empresas e estabelecer regras de classificação para distinguir as firmas estáveis das insolventes.

O conhecimento dos fatores relacionais entre as atividades/metapoliciais e a aparelhagem utilizada por eles em países subdesenvolvidos pode ajudar no processo de modernização e melhorias das condições existentes. O objetivo da Análise Multivariada foi determinar a dependência entre dois tipos de variáveis – atividade → meta e aparelhagem policial.

Os executivos de alto nível estão sempre tomando decisões de risco - Seu sucesso depende do quanto ele conhece o risco de cada negócio. Medidas de variáveis relacionadas com a dimensão desse risco foram identificadas para se obter as características sócio econômica de uma grande quantidade de negócios. O objetivo foi analisar as relações entre a propensão do risco e as características sócio econômicas do negócio.

Quatro diferentes tipos de novos produtos em firmas coreanas, classificadas pela origem e local do fornecimento de produtos estrangeiros, foram observadas a fim de se estudar a padronização das inovações. Medidas feitas com seis variáveis em 42 firmas foram usadas para confirmar as diferenças nos novos padrões. O resultado tem importantes

implicações no gerenciamento de novas firmas em países industrializados. Aqui o objetivo foi verificar se um tipo de variável interna poderia diferenciar significativamente o produto.

d) Educação

Os resultados em testes de aptidão escolar e a performance dos alunos durante o segundo grau são usados frequentemente como indicadores do sucesso acadêmico na universidade. As medidas efetuadas com cinco variáveis em estabelecimentos pré universitários e quatro variáveis sobre performance universitária foram usadas para determinar a associação entre o previsto e o resultado efetivo de escores. O objetivo foi aplicar regras para a classificação de estudantes na universidade.

Atividades atléticas têm sido frequentemente analisadas com o propósito de identificar habilidades primárias em variados eventos. Dados de 8 diferentes decatlos olímpicos serviram de registros para a Análise Multivariada identificar os fatores físicos responsáveis pelos resultados nos 10 esportes do decatlo. Os resultados poderiam ser aparentemente explicados em termos de 4 fatores físicos: velocidade, força no braço, resistência na corrida e força na perna. Aqui o objetivo foi determinar a dependência das variáveis observadas (resultados em pistas e campo) em variáveis menos latentes (fatores físicos).

e) Biologia

No estudo de reprodução de plantas é necessário, após o fim de uma geração, selecionar aquelas plantas que serão os pais da próxima geração. A seleção deve ser feita de tal modo que as gerações sucessivas possam ser melhoradas em um número de características sobre aquelas da geração anterior. A meta da reprodução em plantas é maximizar o ganho genético em um mínimo de tempo. Técnicas multivariadas foram usadas em um programa de

reprodução de feijão para transformar as medições realizadas em variáveis dos rendimentos e das proteínas existentes dentro de um índice de seleção. Os escores neste índice foram então usados para determinar os pais de uma subsequente família de feijão.

Duas espécies de plantas rasteiras são de difícil identificação. Medições realizadas em quatro variáveis dessas plantas rasteiras foram usadas para se obter uma função dos quais os valores permitiram uma separação em dois grupos. Por conseguinte, a função poderia ser usada para classificar uma nova planta, como também identificar se a planta pertence a uma espécie ou a outra.

f) Estudos Ambientais

As concentrações de ar poluente na atmosfera têm sido intensivamente estudadas. Em um estudo, foram realizadas medições diárias sobre a poluição do ar em função de um período de tempo, sendo registrados esses dados em sete tipos de variáveis. O interesse imediato era verificar se o nível de poluição atmosférica era aproximadamente constante durante a semana ou se haviam diferenças consideráveis entre o nível de poluição durante os dias da semana e o fim de semana. Um objetivo secundário era verificar se a massa de dados disponíveis poderia ser resumidos e interpretados de modo rápido. Aqui os objetivos foram testar hipóteses e reduzir dados.

g) Meteorologia

Um estudo foi iniciado para quantificar as relações entre três anéis cronológicos e vários parâmetros climáticos. Era de interesse determinar o tipo de informação climática que cada anel continha e então reconstruir as anomalias climáticas que datam após 1700. Técnicas multivariadas foram usadas para reduzir a quantidade enorme de dados avaliados

para um tamanho manuseável. Os poucos novos dados criados no processo foram analisados e interpretados subsequente de modo muito mais fácil (aqui o objetivo inicial foi reduzir dados).

h) Geologia

A Análise Multivariada foi usada em estudos para classificar tamanho e distribuição de sedimentos em ordem para construir duas funções lineares de 10 tamanhos. Os resultados permitem uma considerável redução no trabalho de laboratório, necessário para diferenciar entre os diferentes tipos de sedimentos (uma vez mais, o objetivo foi reduzir e classificar dados).

i) Psicologia

Um estudo foi realizado para investigar o comportamento de risco. Como parte do estudo, os estudantes foram selecionados aleatoriamente para receber um dos três diferentes tipos de testes. Foram então administrados duas formas paralelas de um teste que penalizavam mais ou menos as respostas incorretas. Os escores das penalidades altas e baixas eram então registradas e ordenadas a fim de analisar as reações dos estudantes em função do risco corrido. O objetivo foi testar hipóteses, ou seja, testar se a direção natural fez diferença em relação ao risco percebido.

j) Esportes

A análise dos recordes em pistas e em campo tem sido freqüentemente foco de atenção nos jogos olímpicos e mundiais. Em um estudo, os recordes nacionais em pistas foram usados para construir a medida ideal de um excelente atleta de uma nação com vistas

nos recordes alcançados. Além disso, a resistência relativa dos atletas de uma nação nas várias distâncias corridas (incluindo a maratona) foi examinada. O resultado para o ranking das nações correspondeu perto da sua performance nos jogos olímpicos. O objetivo foi combinar medidas em várias variáveis dentro de um índice de atleta excelente e examinar a associação entre variáveis representando diversos eventos corridos.

2.4.2 ANÁLISE FATORIAL x ANÁLISE DE COMPONENTES PRINCIPAIS

Existem dois métodos para a análise da estrutura da matriz de covariância ou da matriz de correlação que muito se assemelham mas, que na verdade, têm propósitos diferentes. Um deles é a Análise de Componentes Principais, desenvolvida por Pearson em 1901, através do trabalho intitulado “*On lines and planes of closest fit to a system of points in space*”, páginas 557 a 572, e completada por Hotelling em “*Analysis of a complex of statistical variables into principal components*”, páginas 417 a 441 e 498 a 520, documento datado de 1933. O outro, é a Análise Fatorial, originada a partir de dois trabalhos de Spearman: “*General intelligence objectively determined and measured*”, páginas 201 a 293, de 1904 e “*The Abilities of Man*” de 1926.

Na Análise de Componentes Principais, um conjunto de p variáveis, denotadas por X_1, X_2, \dots, X_p , são transformadas linearmente e ortogonalmente em um número igual de variáveis Y_1, Y_2, \dots, Y_p que têm a propriedade de serem não correlacionadas. Estas variáveis são escolhidas de modo que tenham variâncias máximas.

Se algebricamente, componentes principais são combinações lineares das p variáveis aleatórias, geometricamente, estas combinações lineares representam a seleção de um novo

sistema de coordenadas pela rotação do sistema original de eixos X_1, X_2, \dots, X_p . Os novos eixos representam as direções com máxima variabilidade e produzem uma simples e mais parcimoniosa descrição da estrutura de covariância.

As componentes principais são obtidas a partir dos autovalores e autovetores da matriz de covariância ou de correlação. Os autovalores arranjados em ordem decrescente de magnitude, correspondem as variâncias das variáveis Y , que são componentes principais não padronizados. Frequentemente os primeiros componentes estimam uma grande proporção do total da variância das X variáveis e podem, para certos propósitos, ser usadas para resumirem os dados originais. Entretanto, todas as componentes são necessárias para reproduzir com exatidão os coeficientes de correlação entre as variáveis X . O desenvolvimento da análise de componentes principais não requer a hipótese de que as observações provenham de populações normalmente distribuídas.

A seguir será descrito o modelo básico de componentes principais.

Seja um vetor aleatório $\tilde{X}' = [X_1, X_2, \dots, X_p]$ com matriz de covariância Σ , cujos pares de autovalor-autovetor são $(\lambda_1, \tilde{e}_1), (\lambda_2, \tilde{e}_2), \dots, (\lambda_p, \tilde{e}_p)$ e $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, e as combinações lineares :

$$Y_1 = \tilde{e}_1' X = e_{11}X_1 + e_{21}X_2 + \dots + e_{p1}X_p$$

$$Y_2 = \tilde{e}_2' X = e_{12}X_1 + e_{22}X_2 + \dots + e_{p2}X_p$$

$$\vdots \qquad \qquad \qquad \vdots$$

$$Y_p = \tilde{e}_p' X = e_{1p}X_1 + e_{2p}X_2 + \dots + e_{pp}X_p$$

ou na forma matricial

$$\tilde{Y} = E' \tilde{X}$$

onde E é a matriz de ordem $p \times p$ de autovetores normalizados nas colunas, tal que

$$\text{Var}(Y_i) = e_i' \Sigma e_i = \lambda_i \quad i = 1, 2, \dots, p$$

$$\text{Cov}(Y_i, Y_k) = e_i' \Sigma e_k = 0 \quad i, k = 1, 2, \dots, p, i \neq k$$

O resultado 2.1 seguinte, refere-se ao Modelo de Componentes Principais.

Resultado 2.1:

Seja $B_{p \times p}$ uma matriz positiva definida com autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ associados aos autovetores normalizados e_1, e_2, \dots, e_p . Então

$$\underset{\tilde{x} \neq 0}{\text{máx}} \frac{\tilde{x}' B \tilde{x}}{\tilde{x}' \tilde{x}} = \lambda_1 \quad \text{obtido quando } \tilde{x} = e_1$$

e assim sucessivamente até

$$\underset{\tilde{x} \neq 0}{\text{máx}} \frac{\tilde{x}' B \tilde{x}}{\tilde{x}' \tilde{x}} = \lambda_p \quad \text{obtido quando } \tilde{x} = e_p$$

Além disso,

$$\underset{\tilde{x} \perp e_1, \dots, e_k}{\text{máx}} \frac{\tilde{x}' B \tilde{x}}{\tilde{x}' \tilde{x}} = \lambda_{k+1} \quad \text{obtido quando } \tilde{x} = e_{k+1}, \quad k = 1, 2, \dots, p-1.$$

Prova:

Seja $P_{p \times p}$ uma matriz ortogonal, cujas colunas são os autovetores e_1, e_2, \dots, e_p e seja Λ uma matriz diagonal com autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$ ao longo de sua diagonal. Seja $B^{1/2} = P \Lambda^{1/2} P'$ e $y = P' x$. Conseqüentemente, $x \neq 0$ implica $y \neq 0$. Assim,

$$\frac{\tilde{x}' B \tilde{x}}{\tilde{x}' \tilde{x}} = \frac{\tilde{x} B^{1/2} B^{1/2} \tilde{x}}{\tilde{x}' \underbrace{P P'}_{I_{p \times p}} \tilde{x}} = \frac{\tilde{x}' P \Lambda^{1/2} P' P \Lambda^{1/2} P' \tilde{x}}{y' y} = \frac{y' \Lambda y}{y' y} = \frac{\sum_{i=1}^p \lambda_i y_i^2}{\sum_{i=1}^p y_i^2} \leq \lambda_1 \frac{\sum_{i=1}^p y_i^2}{\sum_{i=1}^p y_i^2} = \lambda_1$$

trocando

$$\tilde{x} = e_{\tilde{1}} \text{ tem-se } \tilde{y} = P' e_{\tilde{1}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ pois } e_{\tilde{k}}' e_{\tilde{1}} = \begin{cases} 1, & k = 1 \\ 0, & k \neq 1 \end{cases}.$$

Para esta escolha de \tilde{x} , $\frac{y' \Lambda y}{y' y} = \frac{\lambda_1}{1} = \lambda_1$, ou $\frac{e_{\tilde{1}}' B e_{\tilde{1}}}{e_{\tilde{1}}' e_{\tilde{1}}} = e_{\tilde{1}}' B e_{\tilde{1}} = \lambda_1$.

Um argumento similar produz a segunda parte.

Agora $\tilde{x} = P' \tilde{y} = y_1 e_{\tilde{1}} + y_2 e_{\tilde{2}} + \dots + y_p e_{\tilde{p}}$, como $\tilde{x} \perp e_{\tilde{1}}, \dots, e_{\tilde{k}}$ implica

$$0 = e_{\tilde{i}}' \tilde{x} = y_1 e_{\tilde{i}}' e_{\tilde{1}} + y_2 e_{\tilde{i}}' e_{\tilde{2}} + \dots + y_p e_{\tilde{i}}' e_{\tilde{p}} = y_i, \quad i \leq k$$

Portanto, para \tilde{x} perpendicular aos primeiros k autovetores $e_{\tilde{i}}$, tem-se:

$$\frac{\tilde{x}' B \tilde{x}}{\tilde{x}' \tilde{x}} = \frac{\sum_{i=k+1}^p \lambda_i y_i^2}{\sum_{i=k+1}^p y_i^2}$$

Tomando $y_{k+1} = 1, y_{k+2} = \dots = y_p = 0$, obtém-se o máximo declarado.

Resultado 2.2:

Seja Σ uma matriz de covariância associada com o vetor aleatório $\underline{X}' = [X_1, X_2, \dots, X_p]$ e $(\lambda_1, \underline{e}_1), (\lambda_2, \underline{e}_2), \dots, (\lambda_p, \underline{e}_p)$ os pares de autovalor-autovetor de Σ , com $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. A i -ésima componente principal é dada por:

$$Y_i = \underline{e}_i' \underline{X} = e_{1i}X_1 + e_{2i}X_2 + \dots + e_{pi}X_p, \quad i = 1, 2, \dots, p$$

Então

$$Var(Y_i) = \underline{e}_i' \Sigma \underline{e}_i = \lambda_i \quad i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_k) = \underline{e}_i' \Sigma \underline{e}_k = 0 \quad i, k = 1, 2, \dots, p, \quad i \neq k$$

Se existirem λ_i iguais, tem-se em consequência Y_i não únicos.

Prova:

Do resultado 2.1, fazendo $B = \Sigma$, tem-se:

$$\underset{\substack{\ell \neq 0 \\ \ell \neq 0}}{\underset{\sim}{\text{máx}}} \frac{\ell' \Sigma \ell}{\ell' \ell} = \lambda_1 \quad (\text{obtido quando } \ell = \underline{e}_1)$$

Se os autovetores são normalizados, $\underline{e}_i' \underline{e}_i = 1$, então

$$\underset{\substack{\ell \neq 0 \\ \ell \neq 0}}{\underset{\sim}{\text{máx}}} \frac{\ell' \Sigma \ell}{\ell' \ell} = \lambda_1 = \frac{\underline{e}_1' \Sigma \underline{e}_1}{\underline{e}_1' \underline{e}_1} = \underline{e}_1' \Sigma \underline{e}_1 = Var(Y_1)$$

Similarmente,

$$\max_{\substack{\ell \perp e_{\sim 1}, e_{\sim 2}, \dots, e_{\sim k} \\ \ell' \ell = 1}} \frac{\ell' \Sigma \ell}{\ell' \ell} = \lambda_{k+1}, \quad k = 1, 2, \dots, p-1$$

Para a escolha de $\ell = e_{\sim k+1}$, com $e_{\sim k+1}' e_{\sim i} = 0$, para $i = 1, 2, \dots, k$ e $k = 1, 2, \dots, p-1$,

$$\frac{e_{\sim k+1}' \Sigma e_{\sim k+1}}{e_{\sim k+1}' e_{\sim k+1}} = e_{\sim k+1}' \Sigma e_{\sim k+1} = \text{Var}(Y_{k+1})$$

No entanto, da definição de autovalor e autovetor, tem-se

$$e_{\sim k+1}' (\Sigma e_{\sim k+1}) = e_{\sim k+1}' (\lambda_{k+1} e_{\sim k+1}) = \lambda_{k+1} (e_{\sim k+1}' e_{\sim k+1}) = \lambda_{k+1}$$

Portanto, $\text{Var}(Y_{k+1}) = \lambda_{k+1}$.

Resta mostrar que $e_{\sim i}$ é perpendicular a $e_{\sim k}$, dado que $\text{Cov}(Y_i, Y_k) = 0$. Os autovetores de Σ são ortogonais se todos os autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$ são distintos. Se os autovalores não são todos distintos, os autovetores correspondentes aos autovalores comuns podem ser escolhidos ortogonais. Então, para quaisquer dois autovetores $e_{\sim i}$ e $e_{\sim k}$, $e_{\sim i}' e_{\sim k} = 0$, $i \neq k$. Desde

que $\Sigma e_{\sim k} = \lambda_k e_{\sim k}$, a pré-multiplicação por $e_{\sim i}'$ leva a

$$\text{Cov}(Y_i, Y_k) = e_{\sim i}' \Sigma e_{\sim k} = e_{\sim i}' \lambda_k e_{\sim k} = \lambda_k e_{\sim i}' e_{\sim k} = 0 \quad \text{para } i \neq k.$$

Assim, completa-se a prova do resultado.

Em contraste com o Método de Componentes Principais, a proposta da Análise Fatorial é avaliar as covariâncias das variáveis observadas em termos de um número muito menor de variáveis hipotéticas não observáveis, chamadas fatores. Com esses fatores, tenta-se explicar a estrutura de covariância das variáveis observadas. Na Análise Fatorial, as variáveis originais são expressas como combinações lineares desses fatores.

Considerando a natureza dos relacionamentos envolvidos, pode-se também perceber a diferença entre os dois tratamentos. Na Análise de Componentes Principais as variáveis Y 's são, por definição, funções lineares das variáveis X 's, e nenhuma hipótese é sugerida. Na Análise Fatorial, por outro lado, assume-se basicamente que as variáveis X 's são dadas por:

$$X_i = \sum_{j=1}^m \ell_{ij} F_j + \varepsilon_i \quad (i = 1, \dots, p)$$

onde F_j é o j -ésimo fator comum, m é o número de fatores comuns considerados, e onde e_i é um resíduo representando as fontes de variação que afetam apenas a variável X_i .

No modelo anterior, assume-se ainda que as p variações aleatórias e_i são independentes entre si e de qualquer um dos m fatores F_j . Os fatores, por sua vez, podem ser correlacionados entre si (caso oblíquo) ou não correlacionados (caso ortogonal). Usualmente a escala de medida é estabelecida em termos de unidades de variância. Por conveniência e sem perda de generalidade, supõe-se que as médias de todas as variáveis são iguais a zero. A variância de e_i , denominada de variância residual ou variância única de X_i , é denotada por ψ_i . O coeficiente ℓ_{ij} é denominado de carregamento de X_i sobre F_j ou alternativamente, como carregamento de F_j em X_i . Na prática, ℓ_{ij} e ψ_i são usualmente parâmetros desconhecidos, que precisam ser estimados de dados experimentais.

O Quadro 2 a seguir mostra as principais diferenças entre a Análise Fatorial e a Análise de Componentes Principais.

QUADRO 2 - CONTRASTES ENTRE ANÁLISE FATORIAL E ANÁLISE DE COMPONENTES PRINCIPAIS

Análise Fatorial	Análise de Componentes Principais
As variáveis são expressas como combinações lineares dos fatores	As componentes principais são expressas como combinações lineares das variáveis
Procura explicar as covariâncias ou correlações	Enfatiza a explicação da variância total
Faz muitas suposições	Essencialmente não requer suposições
Os fatores estão sujeitos a uma rotação arbitrária	As componentes principais são únicas, considerando autovalores distintos da matriz de covariância ou correlação
Alterando o número de fatores, os carregamentos fatoriais estimados se alteram	Qualquer que seja o número de componentes consideradas, os coeficientes são sempre os mesmos

A habilidade em rotacionar e melhorar a interpretação é uma das vantagens da Análise Fatorial sobre a Análise de Componentes Principais. Se a procura e descrição de alguns fatores básicos é o objetivo, a Análise Fatorial pode ser mais útil. Geralmente prefere-se a Análise Fatorial, quando o modelo tem um bom ajuste dos dados e pretende-se interpretar os fatores rotacionados. Por outro lado, quando o objetivo é definir um número menor de variáveis para a entrada em uma outra análise, prefere-se ordinariamente componentes principais, mesmo que isso possa muitas vezes ser realizado com os escores fatoriais.

2.4.3 MODELO FATORIAL ORTOGONAL

Sendo o vetor aleatório observável \tilde{X} , com p componentes, média μ e covariância Σ , o Modelo Fatorial postula que \tilde{X} é linearmente dependente de algumas variáveis aleatórias

não observáveis F_1, F_2, \dots, F_m , chamadas fatores comuns, e p fontes de variação $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$, chamadas erros ou fatores específicos. Em particular, o Modelo de Análise Fatorial é

$$X_1 - \mu_1 = \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1$$

$$X_2 - \mu_2 = \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2$$

⋮

$$X_p - \mu_p = \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p$$

ou em notação matricial,

$$\underset{(px1)}{\tilde{X}} - \underset{(px1)}{\tilde{\mu}} = \underset{(pxm)}{\tilde{L}} \cdot \underset{(mx1)}{\tilde{F}} + \underset{(px1)}{\tilde{\varepsilon}}$$

O coeficiente ℓ_{ij} é chamado de carregamento da i -ésima variável no j -ésimo fator. Assim a matriz L é a matriz de carregamentos. Observa-se ainda que o i -ésimo fator específico ε_i está associado apenas com a i -ésima resposta X_i . Os p desvios $X_1 - \mu_1, X_2 - \mu_2, \dots, X_p - \mu_p$, são expressos em termos de $p+m$ variáveis aleatórias $F_1, F_2, \dots, F_m, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ que não são observáveis. O modelo apresentado difere de um modelo de regressão multivariada, justamente porque as variáveis aleatórias \tilde{F} podem ser observadas no Modelo de Regressão Linear Multivariada, e não podem ser observadas aqui, no Modelo Fatorial Ortogonal.

Assume-se no modelo ainda que:

$$E(\tilde{F}) = \underset{(mx1)}{\tilde{0}}, \text{Cov}(\tilde{F}) = E[\tilde{F}\tilde{F}'] = \underset{(mxm)}{\tilde{I}}$$

A esperança de cada fator é nula e os fatores não são correlacionados. \tilde{F} e $\tilde{\varepsilon}$ são independentes.

$$\text{Cov}(\underline{\varepsilon}, \underline{F}) = E(\underline{\varepsilon} \underline{F}') = \underset{(pxm)}{0}$$

Com essas considerações, pode-se organizar um Modelo Fatorial Ortogonal com m fatores comuns do seguinte modo:

$$\underset{(px1)}{X} = \underset{(px1)}{\mu} + \underset{(pxm)}{L} \cdot \underset{(mx1)}{F} + \underset{(px1)}{\varepsilon}$$

com $m < p$ e, ainda

μ_i = média da variável i ; $i = 1, 2, 3, \dots, p$

ε_i = i -ésimo fator específico

F_j = j -ésimo fator comum; $j = 1, 2, 3, \dots, m$

ℓ_{ij} = carregamento da i -ésima variável no j -ésimo fator

Os não observáveis vetores aleatórios \underline{F} e $\underline{\varepsilon}$ satisfazem as propriedades

- \underline{F} e $\underline{\varepsilon}$ são independentes
- $E(\underline{F}) = \underline{0}$, $\text{Cov}(\underline{F}) = I$
- $E(\underline{\varepsilon}) = \underline{0}$, $\text{Cov}(\underline{\varepsilon}) = \Psi$, onde Ψ é uma matriz diagonal com a variância específica ψ_i na diagonal principal

Com os resultados 2.3 e 2.4, pode-se provar que, com as suposições do modelo, tem-se:

$$\Sigma = \text{Cov}(\underline{X}) = LL' + \Psi$$

e

$$\text{Cov}(\underline{X}, \underline{F}) = L$$

Resultado 2.3 :

Com as suposições do Modelo Fatorial Ortogonal, tem-se que

$$\Sigma = Cov(\underline{X}) = LL' + \Psi$$

onde Σ é a matriz de covariância, L é matriz de carregamentos fatoriais e Ψ é a matriz de variâncias específicas.

Prova:

Da definição de covariância de uma variável aleatória, tem-se

$$\Sigma = Cov(\underline{X}) = E(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'$$

isolando μ no Modelo Fatorial Ortogonal e substituindo na equação anterior, vem

$$\Sigma = Cov(\underline{X}) = E(L\underline{F} + \underline{\varepsilon})(L\underline{F} + \underline{\varepsilon})' = E[L\underline{F}(\underline{L}\underline{F})' + \underline{\varepsilon}(\underline{L}\underline{F})' + L\underline{F}\underline{\varepsilon}' + \underline{\varepsilon}\underline{\varepsilon}']$$

$$\Sigma = Cov(\underline{X}) = E[L\underline{F}\underline{F}'L' + \underline{\varepsilon}\underline{F}'L' + L\underline{F}\underline{\varepsilon}' + \underline{\varepsilon}\underline{\varepsilon}']$$

usando propriedades da esperança

$$\Sigma = Cov(\underline{X}) = E(L\underline{F}\underline{F}'L') + E(\underline{\varepsilon}\underline{F}'L') + E(L\underline{F}\underline{\varepsilon}') + E(\underline{\varepsilon}\underline{\varepsilon}')$$

$$\Sigma = Cov(\underline{X}) = LE(\underline{F}\underline{F}')L' + E(\underline{\varepsilon}\underline{F}')L' + LE(\underline{F}\underline{\varepsilon}') + E(\underline{\varepsilon}\underline{\varepsilon}')$$

das suposições do Modelo Fatorial Ortogonal, pode-se escrever

$$\Sigma = Cov(\underline{X}) = LIL' + 0L' + L0 + \Psi$$

$$\Sigma = Cov(\underline{X}) = LL' + \Psi$$

Resultado 2.4 :

Com as suposições do Modelo Fatorial Ortogonal, tem-se que

$$\text{Cov}(\underline{X}, \underline{F}) = L$$

Prova:

Das propriedades de covariância entre duas variáveis, pode-se escrever:

$$\text{Cov}(\underline{X}, \underline{F}) = E(\underline{X} - \underline{\mu})(\underline{F})'$$

Isolando $\underline{X} - \underline{\mu}$ no Modelo Fatorial Ortogonal e substituindo na equação anterior, vem

$$\text{Cov}(\underline{X}, \underline{F}) = E(L\underline{F} + \underline{\varepsilon})(\underline{F})' = E(L\underline{F}\underline{F}' + \underline{\varepsilon}\underline{F}')$$

usando propriedades da esperança

$$\text{Cov}(\underline{X}, \underline{F}) = LE(\underline{F}\underline{F}') + E(\underline{\varepsilon}\underline{F}')$$

das suposições do Modelo Fatorial Ortogonal, pode-se escrever

$$\text{Cov}(\underline{X}, \underline{F}) = LI + 0$$

$$\text{Cov}(\underline{X}, \underline{F}) = L$$

Desta forma, com os resultados anteriores, pode-se montar uma Estrutura de Covariância para o Modelo Fatorial Ortogonal :

$$1. \text{Cov}(\underline{X}) = LL' + \Psi$$

ou

$$\text{Var}(X_i) = l_{i1}^2 + \dots + l_{im}^2 + \psi_i$$

$$\text{Cov}(X_i, X_k) = l_{i1}l_{k1} + \dots + l_{im}l_{km}$$

$$2. \text{Cov}(\underline{X}, \underline{F}) = L$$

ou

$$\text{Cov}(X_i, F_j) = l_{ij}$$

O modelo anterior é adequado quando se assume um relacionamento linear entre as respostas \underline{X} e os fatores \underline{F} . A suposição de linearidade é inerente ao tradicional Modelo Fatorial.

A porção da variância de uma variável X_i , advinda como contribuição dos m fatores comuns, é denominada comunalidade i , e denotada por h_i^2 . A porção da variância desta variável atribuída ao fator específico ε_i é chamada de imparidade ou, mais freqüentemente, variância específica que é denotada por ψ_i .

Portanto, a variância de qualquer variável do modelo é composta de duas partes : a comunalidade e a variância específica. Em particular, para a variável i , tem-se:

$$\text{Var}(X_i) = (\text{comunalidade } i) + (\text{variância específica } i)$$

Como a i -ésima comunalidade é a soma dos quadrados dos carregamentos da i -ésima variável nos m fatores comuns, ou seja,

$$h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2$$

e $\text{Var}(X_i) = \sigma_{ii}$, então a relação resumida é

$$\sigma_i^2 = \sigma_{ii} = h_i^2 + \psi_i, \quad i = 1, 2, \dots, p$$

O Modelo Fatorial assume que $p + \binom{p}{2} = p + \frac{p(p-1)}{2} = \frac{p(p+1)}{2}$ variâncias e covariâncias de \underline{X} podem ser reproduzidas pelos $p \cdot m$ carregamentos fatoriais ℓ_{ij} e as p variâncias específicas ψ_i . Quando $m = p$, qualquer matriz de covariância Σ pode ser

reproduzida exatamente como LL' , então Ψ pode ser a matriz nula. Entretanto, a análise mais usada é aquela em que m é menor do que p . Neste caso, o Modelo Fatorial produz uma simples explicação da covariância em \tilde{X} com menos parâmetros em relação aos $\frac{p(p+1)}{2}$ parâmetros de Σ .

Em alguns casos, quando se considera um Modelo Ortogonal Fatorial em que m é igual a 1, ou seja, tem-se apenas um fator, pode-se também obter uma única solução numérica. Entretanto, esta solução numérica pode não ser consistente com a interpretação estatística dos coeficientes, gerando, portanto, uma solução imprópria. A inconsistência da solução em relação aos coeficientes acontece quando no cálculo dos carregamentos, aparecem no modelo valores para variâncias maiores que 1 em valor absoluto, ou até mesmo valores negativos de variâncias específicas.

Para $m > 1$, existe sempre uma ambigüidade inerente associada com o Modelo Fatorial.

Como ilustração, seja uma matriz T de ordem $m \times m$ ortogonal, ou seja, $TT' = T'T = I$. Assim, o Modelo Fatorial Ortogonal poderia ser escrito:

$$\tilde{X} - \tilde{\mu} = L\tilde{F} + \tilde{\varepsilon}$$

$$\tilde{X} - \tilde{\mu} = LTT' \tilde{F} + \tilde{\varepsilon}$$

$$\tilde{X} - \tilde{\mu} = (LT)(T' \tilde{F}) + \tilde{\varepsilon}$$

fazendo $LT = L^*$ e $T' \tilde{F} = \tilde{F}^*$, tem-se

$$\tilde{X} - \tilde{\mu} = L^* \tilde{F}^* + \tilde{\varepsilon}$$

E assim, preservam-se as suposições, mesmo com a transformação ortogonal.

$$E(\underline{F}^*) = E(T' \underline{F}) = T' \cdot E(\underline{F}) = T' \cdot 0 = 0$$

e

$$Cov(\underline{F}^*) = Cov(T' \underline{F}) = T \cdot Cov(\underline{F}) \cdot T' = T \cdot I \cdot T' = T \cdot T' = I$$

No modelo, não é possível distinguir os carregamentos de L dos carregamentos de L^* . Os carregamentos fatoriais de F e F^* têm as mesmas propriedades estatísticas. Mesmo que L e L^* tenham diferentes carregamentos, ambos são gerados pela mesma matriz de covariância Σ . Assim, não há diferença para o modelo entre $L^* = LT$ e L , pois ambas têm a mesma representação. As comunalidades que são os elementos diagonais da matriz $LL' = (L^*)(L^*)'$ são invariantes pela escolha de T . Logo, pode-se escrever:

$$\Sigma = LL' + \Psi = LTT'L' + \Psi = (L^*)(L^*)' + \Psi$$

Esta ambigüidade produz a possibilidade de se estabelecer uma rotação fatorial. Da Álgebra Matricial, sabe-se que uma transformação ortogonal corresponde a uma rígida rotação (ou reflexão) dos eixos coordenados. Mantendo inalteradas as comunalidades e as variâncias específicas, pode-se encontrar uma matriz ortogonal T de rotação que estime a estrutura de covariância da variável \underline{X} de uma forma mais simples.

2.4.4 SINGULARIDADE DOS PARÂMETROS

Muitas vezes a unicidade de alguns parâmetros está sujeito a restrições impostas a outros parâmetros. Dependendo das condições, pode-se observar para um parâmetro, uma infinidade de possibilidades.

Seja H_m a hipótese de que, para um determinado valor de m (número de fatores), uma dada matriz de covariância Σ pode ser expressa na forma

$$\Sigma = LL' + \Psi. \quad (I)$$

Se uma satisfatória matriz Ψ pode ser encontrada, L pode ser definida unicamente em termos de Σ e Ψ . Considera-se, portanto, que existe uma satisfatória matriz Ψ , se ela é única.

A condição, para $m > 1$, que $L'\Psi^{-1}L$ pode ser diagonal tem imposto, como efeito, $\frac{1}{2}m(m-1)$ restrições sobre os parâmetros. Assim, o número de parâmetros livres em Ψ e L é

$$p + pm - \frac{1}{2}m(m-1)$$

Igualando os correspondentes elementos matriciais de ambos os membros de (I), obtém-se $\frac{1}{2}p(p+1)$ distintas equações. Sendo s a diferença entre o número de equações e o número de parâmetros livres, então

$$s = \frac{1}{2}p(p+1) - p - pm + \frac{1}{2}m(m-1)$$

$$s = \frac{1}{2}[(p-m)^2 - (p+m)]$$

Se $s = 0$, tem-se muitas equações com parâmetros livres, então é de se esperar que, em geral, a hipóteses H_m seja trivialmente verdadeira e Ψ seja unicamente determinada. Se $s < 0$, existem menos equações que parâmetros livres, então espera-se que H_m seja também trivialmente verdadeira, mas com uma infinidade de escolhas para Ψ e L . Por outro lado, se $s > 0$, existem mais equações do que parâmetros livres. Neste caso, H_m não é trivial. A hipótese é verdadeira apenas se existem restrições sobre os elementos de Σ .

Como ilustração, são considerados 4 exemplos de algumas situações no cálculo dos elementos de L .

a) Seja $\Sigma = \begin{bmatrix} 1 & 0,56 & 0,40 \\ 0,56 & 1 & 0,35 \\ 0,40 & 0,35 & 1 \end{bmatrix}$. Para encontrar os carregamentos de L , resolve-se as

equações

$$\ell_1 \ell_2 = 0,56 \quad \ell_1 \ell_3 = 0,40 \quad \ell_2 \ell_3 = 0,35$$

A solução leva a $\ell_1 = 0,8$, $\ell_2 = 0,7$ e $\ell_3 = 0,5$. As variâncias específicas na diagonal de Ψ seriam 0,36, 0,51 e 0,75, respectivamente. A solução neste caso é única.

b) Seja $\Sigma = \begin{bmatrix} 1 & 0,84 & 0,60 \\ 0,84 & 1 & 0,35 \\ 0,60 & 0,35 & 1 \end{bmatrix}$. Resolvendo obtém-se $\ell_1 = 1,2$, $\ell_2 = 0,7$ e $\ell_3 = 0,5$.

Desta forma, a variância específica ψ_1 seria negativa, ou seja,

$$\psi_1 = 1 - \ell_1^2 = -0,44$$

Neste caso, não existe uma solução aceitável com $m = 1$, uma vez que existe em Ψ , elementos não positivos. A matriz $\Sigma^* - I$, com $\Sigma^* = \Psi^{-1/2} \Sigma \Psi^{-1/2}$, tem um autovalor positivo e os outros dois nulos. Caso se considere dois fatores, existe uma infinidade de soluções.

c) Como um exemplo adicional de uma solução inconsistente, pode-se supor que os três elementos fora da diagonal principal de Σ sejam negativos. Na solução de ℓ_1 , ℓ_2 e ℓ_3 encontram-se números imaginários. A dificuldade aqui, é que, mesmo que Ψ tenha elementos positivos, o único autovalor não nulo de $\Sigma^* - I$ é negativo.

d) Pode-se supor, também, que Σ tenha elementos diagonais unitários e elementos não diagonais nulos, exceto σ_{12} , que satisfaz $0 < \sigma_{12}^2 < 1$. Desde que ℓ_1 e ℓ_2 satisfaçam as condições

$$\ell_1 \ell_2 = \sigma_{12} \quad \ell_1^2 < 1 \quad \ell_2^2 < 1$$

existe uma infinidade de soluções.

Este último caso viola a condição necessária para a unicidade de Ψ , que impõe que para qualquer valor de m , cada coluna de L deva ter os menores três elementos não nulos, mesmo que os fatores sejam rotacionados.

2.4.5 NÚMERO DE FATORES COMUNS DA ANÁLISE FATORIAL

Um item muito importante numa Análise Fatorial é a escolha do número de fatores comuns. A adequabilidade do modelo e a interpretação dos resultados dependem significativamente do número de fatores adotado. Dentre os diversos critérios que são normalmente utilizados para definir o número m , na presente dissertação, cinco deles serão abordados. Os quatro primeiros são largamente aplicados no Método de Componentes Principais para a estimação dos carregamentos fatoriais da matriz L .

CRITÉRIO 1:

O número de fatores m é escolhido de acordo com o percentual da variância explicada.

A proporção do total da variância amostral atribuída ao j -ésimo fator é

$$\frac{\hat{\ell}_{1j}^2 + \hat{\ell}_{2j}^2 + \dots + \hat{\ell}_{pj}^2}{tr(S)} = \frac{\lambda_j}{tr(S)}$$

quando a análise é feita a partir de S , o estimador de Σ . A correspondente proporção quando a análise é feita a partir de R , o estimador de ρ é

$$\frac{\hat{\ell}_{1j}^2 + \hat{\ell}_{2j}^2 + \dots + \hat{\ell}_{pj}^2}{p} = \frac{\lambda_j}{p}$$

onde p é o número de variáveis. Assim, a contribuição de todos os m fatores para $\text{tr}(S)$ ou p é a soma dos quadrados de todos os elementos de \hat{L} , ou seja $\sum_{i=1}^p \sum_{j=1}^m \hat{\ell}_{ij}^2$. Para o Método de Componentes Principais, esta soma é também a soma das comunalidades ou a soma dos primeiros autovalores λ :

$$\sum_{i=1}^p \sum_{j=1}^m \hat{\ell}_{ij}^2 = \sum_{i=1}^p \hat{h}_i^2 = \sum_{j=1}^m \lambda_j$$

Desta forma, a escolha de m é feita objetivando uma explicação suficiente da variância total amostral, relativamente a $\text{tr}(S)$, para matriz de covariância amostral ou p , para a matriz de correlação amostral.

O critério 1 pode também ser estendido para o Método do Fator Principal, no qual estimativas iniciais das comunalidades são usadas para a forma $S - \hat{\Psi}$ ou $R - \hat{\Psi}$. Entretanto, $S - \hat{\Psi}$ ou $R - \hat{\Psi}$ podem apresentar autovalores negativos, pois foram retiradas as variâncias específicas. Como os valores de m variam de 1 a p , a proporção cumulativa dos autovalores

$\frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^p \lambda_j}$ pode ser maior que 1, somando-se autovalores negativos. Conseqüentemente, uma

porcentagem, como por exemplo 80%, será alcançada para um menor valor de m , tanto para S quanto para R . Para que isso não ocorra, uma estratégia é escolher m igual ao número de autovalores para os quais a primeira porcentagem exceda 100%. Assim, não se corre tanto o risco de um percentual de explicação insuficiente da variância total amostral.

No método iterativo do Fator Principal, m é determinado antes das iterações, e $\text{tr}(S - \hat{\Psi})$, depois das iterações. Mesmo assim, o valor de m pode ser escolhido a partir autovalores como no Método de Componentes Principais.

CRITÉRIO 2:

O número de fatores m é igual ao número de autovalores positivos de S , caso a matriz de covariância amostral é fatorada, ou igual ao número de autovalores maiores que 1 (um) de R , no caso da matriz de correlação amostral ser fatorada.

Este critério é fundamentalmente empírico e está presente em vários programas estatísticos computacionais. Experiências revelam que, na prática, funciona muito bem fornecendo um bom ajuste.

CRITÉRIO 3:

O número de fatores m é igual ao número de autovalores maiores que a média dos autovalores.

A média dos autovalores para R é 1 e para S é $\frac{\sum_{i=1}^p \lambda_i}{p}$. Assim como o critério anterior,

é fundamentalmente heurístico e está presente em vários programas estatísticos computacionais. Foi proposto a primeira vez por Guttman em 1954 em seu trabalho intitulado “*Some necessary conditions for common factor analysis*”, *Psychometrika* 19, p.149-161. Kayser, em trabalhos datados de 1960, “*The application of electronic computers to factor analysis*”, *Educational and Psychological Measurement* 20, p.141-151; e 1974, “*An index of factorial simplicity*”, *Psychometrika* 39, p.31-36, apresentou boas razões para se confirmar a confiabilidade do critério, mostrando que funciona muito bem na prática.

CRITÉRIO 4:

O número de fatores m é escolhido a partir do Teste *Scree*.

O Teste de *Scree* foi assim nomeado por Catell em 1966, em seu trabalho “*The Scree test for the number of factors*”, *Multivariate Behavioral Research* 1, p.245-276, após a denominação *Scree* geológico que se refere aos escombros no fundo de um precipício rochoso.

Este teste gráfico está baseado nos pares ordenados (i, λ_i) , $i = 1, 2, \dots, p$ de S ou R . O gráfico de pontos do IR^2 apresenta o número de autovalores no eixo horizontal e a medida correspondente dos autovalores no eixo vertical, conforme Figura 2.

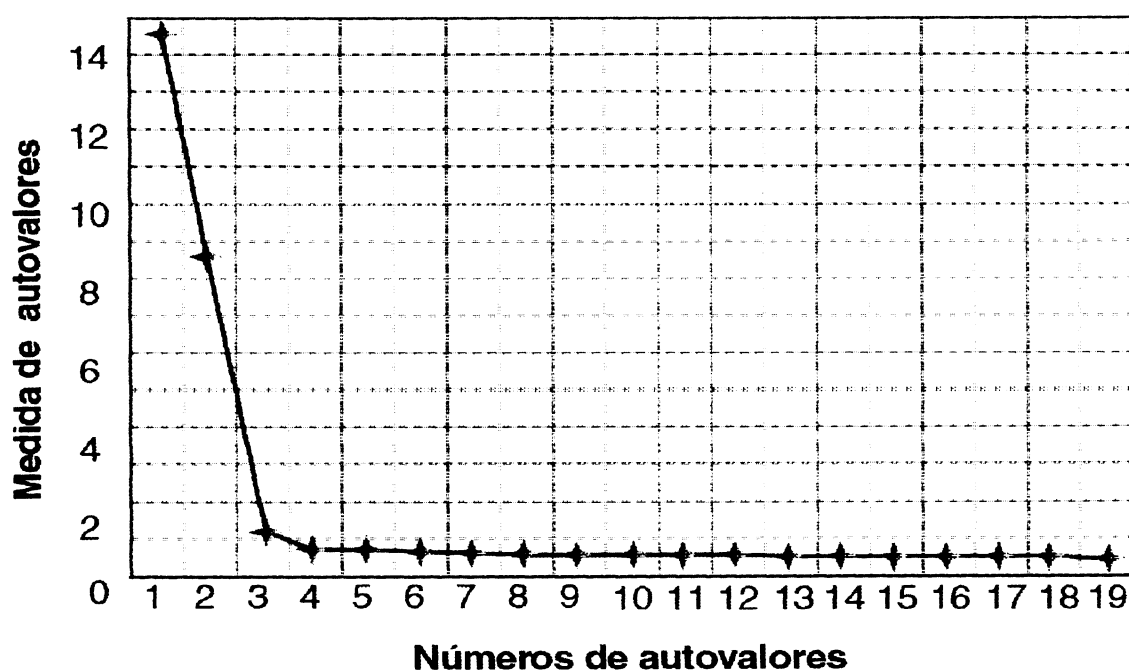


FIGURA 2 - TESTE SCREE

A partir do gráfico, fica fácil distinguir os maiores autovalores dos menores. Desta forma, segundo o Teste *Scree*, escolhe-se o número de fatores m igual ao número de autovalores que estão na primeira linha íngreme. Descarta-se portanto, os menores autovalores que estão na linha mais próxima do eixo horizontal. É importante ressaltar que

um modelo que apresente um número muito grande de pequenos autovalores, segundo este critério, pode apresentar baixa precisão.

Baseado num experimento onde o número de fatores era conhecido, Cattell em 1966, afirmou que a invariabilidade *scree* inicia no k -ésimo autovalor quando k é o número verdadeiro de fatores.

Em um trabalho de 1970, intitulado “*A second generation little jiffy*”, Psychometrika 35, p.401-415, Kayser afirmou que a análise do número de fatores é subjetiva quando existe mais de uma linha principal no gráfico.

A precisão do Teste *Scree* na escolha do número correto de fatores tem sido investigada em diversos estudos e por diversos autores. Em 1982, Hakstian e outros, em “*The behavior of numbers of factors rules simulated data*”, Multivariate Behavioral Research 17, p.193-219, comparou o critério 4 com o critério 3, e concluiu que ambos têm precisão semelhante, quando o número de variáveis é maior do que 250, com variáveis moderadamente inter-relacionadas. Quando as correlações são pequenas, mais componentes são necessárias e ambos os critérios têm baixa precisão. Entretanto, o critério do autovalor médio (3) tem uma performance superior em relação ao critério do Teste *Scree*.

CRITÉRIO 5:

O número de fatores m é escolhido de acordo com o teste de hipóteses:

$$H_0 : \underset{(pxp)}{\Sigma} = \underset{(pxm)}{L} \underset{(m \times p)}{L'} + \underset{(pxp)}{\Psi} ;$$

$H_1 : \Sigma$ é qualquer outra matriz positiva definida.

A suposição do Modelo Fatorial Ortogonal de que a amostra provém de uma população normalmente distribuída leva a um teste de hipóteses para verificar o ajuste e

suficiência do modelo aos dados, dependendo do valor de m considerado. Este teste está associado ao Método da Máxima Verossimilhança para estimar L e Ψ .

Quando Σ não tem uma forma especial, de acordo com o resultado 2.10 e $\hat{\Sigma} = ((n-1)/n)S = S_n$, a função de verossimilhança é proporcional a

$$|S_n|^{-n/2} e^{-np/2} \quad (\text{I})$$

Sobre H_0 , Σ é restrita a forma da hipótese H_0 . Neste caso, considerando \hat{L} e $\hat{\Psi}$ estimadores, respectivamente de L e Ψ , o máximo da função verossimilhança é proporcional a

$$|\hat{\Sigma}|^{-np/2} \exp\left(-\frac{1}{2} \text{tr}\left[\hat{\Sigma}^{-1}\left(\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})'\right)\right]\right) = |\hat{L}\hat{L}' + \hat{\Psi}|^{-np/2} \exp\left(-\frac{1}{2} \text{tr}\left[\left(\hat{L}\hat{L}' + \hat{\Psi}\right)^{-1} S_n\right]\right) \quad (\text{II})$$

Se θ é o vetor composto por todos os parâmetros populacionais desconhecidos, $L(\theta)$ a função de verossimilhança obtida avaliando a densidade conjunta de X_1, X_2, \dots, X_n pelos seus valores observados x_1, x_2, \dots, x_n , Θ um espaço v -dimensional e Θ_0 um espaço v_0 -dimensional, com $\Theta_0 \subset \Theta$, então para uma amostra significativamente grande,

$$-2 \ln \Lambda = -2 \ln \left(\frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} \right)$$

é aproximadamente uma variável aleatória $\chi^2_{v-v_0}$, na qual $v-v_0$ é o número de graus de liberdade. Desta forma, usando (I) e (II), a razão estatística de verossimilhança para testar H_0 é

$$-2 \ln \Lambda = -2 \ln \left[\frac{\text{verossimilhança maximizada sobre } H_0}{\text{verossimilhança maximizada}} \right]$$

ou

$$-2 \ln \Lambda = -2 \ln \left(\frac{|\hat{\Sigma}|}{|S_n|} \right)^{-n/2} + n \left[\text{tr} \left(\hat{\Sigma}^{-1} S_n \right) - p \right]$$

com $v - v_0 = \frac{1}{2} p(p+1) - [p(m+1) - \frac{1}{2} m(m-1)] = \frac{1}{2} [(p-m)^2 - p - m]$ graus de liberdade.

Se $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ é uma amostra aleatória de população normal, as estimativas de máxima verossimilhança \hat{L} e $\hat{\Psi}$ que são obtidas pela maximização de

$$L(\mu, \Sigma) = (2\pi)^{-\frac{(n-1)p}{2}} \cdot |\Sigma|^{-\frac{(n-1)}{2}} \cdot e^{-\left(\frac{1}{2}\right) \text{tr} \left[\Sigma^{-1} \left(\sum_{j=1}^n (\tilde{x}_j - \bar{x})(\tilde{x}_j - \bar{x})' \right) \right]} \cdot (2\pi)^{-\frac{p}{2}} \cdot |\Sigma|^{-\frac{1}{2}} \cdot e^{-\left(\frac{n}{2}\right) (\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu)}$$

sujeitas a condição de singularidade que é

$$L' \Psi^{-1} L = \Delta \text{ ser uma matriz diagonal,}$$

satisfazem

$$\begin{pmatrix} \hat{\Psi}^{-1/2} & \hat{S}_n \hat{\Psi}^{-1/2} \end{pmatrix} \begin{pmatrix} \hat{\Psi}^{-1/2} \hat{L} \end{pmatrix} = \begin{pmatrix} \hat{\Psi}^{-1/2} \hat{L} \end{pmatrix} \begin{pmatrix} I + \hat{\Delta} \end{pmatrix}$$

Assim sendo, a j -ésima coluna de $\hat{\Psi}^{-1/2} \hat{L}$ é o vetor (não normalizado) de $\hat{\Psi}^{-1/2} \hat{S}_n \hat{\Psi}^{-1/2}$ correspondendo ao autovalor $1 + \hat{\Delta}_i$, onde

$$S_n = \frac{1}{n} \sum_{j=1}^n (\tilde{x}_j - \bar{x})(\tilde{x}_j - \bar{x})' = \left(\frac{n-1}{n} \right) S \text{ e } \hat{\Delta}_1 \geq \hat{\Delta}_2 \geq \dots \geq \hat{\Delta}_m.$$

Além disso, na convergência

$$\hat{\psi}_i \text{ é o } i\text{-ésimo elemento diagonal de } S_n - \hat{L} \hat{L}'$$

e

$$\text{tr} \left(\hat{\Sigma}^{-1} S_n \right) = p$$

Desta forma, $\text{tr}\left(\hat{\Sigma}^{-1} S_n\right) - p = 0$ produz $\hat{\Sigma} = \hat{L}\hat{L}' + \hat{\Psi}$, que é a estimativa de máxima

verossimilhança de $\Sigma = LL' + \Psi$. Assim, tem-se

$$-2 \ln \Lambda = -2 \left(-\frac{n}{2}\right) \ln \left(\frac{|\hat{\Sigma}|}{|S_n|} \right) + n[0]$$

ou

$$-2 \ln \Lambda = n \ln \left(\frac{|\hat{\Sigma}|}{|S_n|} \right)$$

BARTLETT (1954, p. 296-298) mostrou que a aproximação Qui-quadrado para a distribuição amostral $-2 \ln \Lambda$ pode ser melhorada pela troca de n na última equação por $(n - 1 - (2p + 4m + 5)/6)$. Desta forma, usando a correção de Bartlett, a hipótese H_0 é rejeitada com um nível de significância α se

$$\left(n - 1 - \frac{(2p + 4m + 5)}{6} \right) \ln \left(\frac{|\hat{L}\hat{L}' + \hat{\Psi}|}{|S_n|} \right) > \chi^2_{[(p-m)^2 - p - m] / 2} (\alpha) \quad (\text{III})$$

desde que n e $n - p$ sejam grandes. Além disso, se o número de graus de liberdade é positivo, então m deve satisfazer

$$m < \frac{1}{2}(2p + 1 - \sqrt{8p + 1}) \quad (\text{IV})$$

para que o teste (III) seja aplicado .

Na prática, o procedimento descrito não pode ser visto apenas como um teste de hipóteses. O problema principal é se o número de fatores escolhido apresenta realmente um bom ajuste para os dados. Assim, pode-se utilizar um procedimento iterativo para a determinação de um valor satisfatório para m . Começando com um pequeno valor m_1 , que

pode ser 1, ou mesmo 0, os parâmetros são estimados com $m = m_1$. Tendo encontrado a solução própria, aplica-se o teste de hipóteses. Se o resultado não for significativo para o nível escolhido, a hipótese nula é aceita com $m = m_1$. Entretanto, se o resultado for significativo, tenta-se com $m = m_1 + 1$. O procedimento continua, com m sendo incrementado em 1 a cada passo, até que a hipótese nula seja aceita para algum m ou rejeitada para $m = m_2$, onde m_2 é o maior valor de m , para o qual (IV) seja satisfeito, ou seja, os graus de liberdade sejam positivos.

Por outro lado, este procedimento pode também ser questionado em relação ao nível de significância dos testes, pelo fato de que a sucessão de hipóteses testadas não leva em consideração as rejeições das hipóteses precedentes. Mesmo assim, na prática, o uso do nível de significância de χ^2 parece não causar sérios erros. Considerando que o modelo está correto com m^* como verdadeiro número de fatores, existe uma pequena probabilidade de que o valor determinado pela seqüência de hipóteses seja menor do que m^* . A probabilidade de que o número de fatores m seja maior do que m^* não pode exceder a probabilidade nominal de significância P . Geralmente essa probabilidade será um pouco menor do que P .

Se a hipótese nula em uma determinada iteração é aceita, isto não quer dizer que o número de fatores determinado pela iteração seja o verdadeiro valor. Desde que fatores adicionais não possam ser diferenciados de erros experimentais, tudo isto apenas significa que não há nada ajustando os fatores adicionais aos dados. Por outro lado, se o número de fatores m é muito grande, isto pode dificultar a interpretação de todos os fatores, mesmo depois de uma transformação ou rotação.

Resultado 2.5 :

Toda matriz de covariância é positiva semi-definida.

Prova:

Sendo a matriz de covariância do vetor aleatório Y com p componentes

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix},$$

pela decomposição espectral, para $\forall X \neq 0$, tem-se $X' \Sigma X = \sigma_{ii}$.

Assim,

$$\sigma_{11} = E(Y_1 - \mu_1)^2 \geq 0$$

$$\sigma_{22} = E(Y_2 - \mu_2)^2 \geq 0$$

.....

$$\sigma_{pp} = E(Y_p - \mu_p)^2 \geq 0$$

Então $\sigma_{ii} \geq 0 \quad \forall i$. Portanto Σ é positiva semi-definida.

Resultado 2.6 :

Todo autovalor de uma matriz quadrada positiva definida é positivo.

Prova:

Seja λ um autovalor de uma matriz definida positiva A e \underline{e} o autovetor correspondente. Para este autovetor \underline{e} tem-se $\langle A\underline{e}, \underline{e} \rangle > 0$. Portanto

$$\langle A\underline{e}, \underline{e} \rangle = \langle \lambda \underline{e}, \underline{e} \rangle = \lambda \langle \underline{e}, \underline{e} \rangle > 0.$$

Uma vez que \underline{e} é um autovetor não nulo, pela propriedade dos produtos internos, $\langle \underline{e}, \underline{e} \rangle$ é real e positivo. Desta forma, dividindo $\lambda \langle \underline{e}, \underline{e} \rangle > 0$ por $\langle \underline{e}, \underline{e} \rangle$ tem-se

$$\lambda \langle \underline{e}, \underline{e} \rangle > 0 \Rightarrow \frac{\lambda \langle \underline{e}, \underline{e} \rangle}{\langle \underline{e}, \underline{e} \rangle} > 0 \Rightarrow \lambda > 0$$

Assim, todo autovalor de uma matriz positiva definida é positivo.

Resultado 2.7 :

Se Σ é positiva definida então Σ^{-1} existe e $\Sigma \underline{e} = \lambda \underline{e}$ implica $\Sigma^{-1} \underline{e} = \left(\frac{1}{\lambda}\right) \underline{e}$ onde (λ, \underline{e}) é um par de autovalor-autovetor de Σ correspondendo ao par $(1/\lambda, \underline{e})$ de Σ^{-1} . Além disso, Σ^{-1} também é positiva definida.

Prova:

Para Σ positiva definida e $\underline{e} \neq 0$ um autovetor, tem-se $0 < \underline{e}' \Sigma \underline{e} = \underline{e} \Sigma \underline{e} = \underline{e}' (\Sigma \underline{e}) = \underline{e}' (\lambda \underline{e}) = \lambda \underline{e}' \underline{e} = \lambda$. Além disso, $\underline{e} = \Sigma^{-1} (\Sigma \underline{e}) = \Sigma^{-1} (\lambda \underline{e})$ ou

$\underline{e} = \lambda \Sigma^{-1} \underline{e}$ e dividindo por $\lambda > 0$ leva a $\Sigma^{-1} \underline{e} = (1/\lambda, \underline{e})$. Assim $(1/\lambda, \underline{e})$ é um par de autovalor-autovetor de Σ^{-1} . Para qualquer vetor \underline{x} de dimensão p , tem-se

$$\underline{x}' \Sigma^{-1} \underline{x} = \underline{x}' \left(\sum_{i=1}^p \left(\frac{1}{\lambda_i} \right) \underline{e}_i \underline{e}_i' \right) \underline{x} = \sum_{i=1}^p \left(\frac{1}{\lambda_i} \right) (\underline{x}' \underline{e}_i)^2 \geq 0$$

desde que cada termo $\lambda_i^{-1} (\underline{x}' \underline{e}_i)^2$ não seja negativo. Além disso, $\underline{x}' \underline{e}_i = 0$ para todo i apenas se

$\underline{x} = \underline{0}$. Então $\underline{x} \neq \underline{0}$ implica $\sum_{i=1}^p (1/\lambda_i) (\underline{x}' \underline{e}_i)^2 > 0$ e Σ^{-1} é positiva definida.

Resultado 2.8:

O determinante da matriz quadrada pode ser expresso como produto dos seus autovalores.

Prova:

Sendo A uma matriz quadrada de ordem $p \times p$, P a matriz dos p autovetores coluna normalizados $[\underline{e}_1, \underline{e}_2, \dots, \underline{e}_p]$ de A e Λ a matriz diagonal dos autovalores de A , pode-se

escrever, pela decomposição espectral, $A = \sum_{i=1}^p \lambda_i \underline{e}_i \underline{e}_i' = P \Lambda P'$, onde $P'P = PP' = I$, I é a

matriz identidade de ordem p . Então, tomando o determinante a ambos os membros da equação $A = P \Lambda P'$, tem-se

$$|A| = |P \Lambda P'|$$

usando propriedades dos determinantes, pode-se escrever:

$$|A| = |P| \cdot |\Lambda| \cdot |P^T|$$

mas

$$|P| \cdot |P^T| = |I| = 1$$

portanto

$$|A| = |P| \cdot |P^T| \cdot |\Lambda|$$

fica

$$|A| = 1 \cdot |\Lambda|$$

$$|A| = \begin{vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_p \end{vmatrix}$$

assim

$$|A| = \lambda_1 \lambda_2 \dots \lambda_p$$

Resultado 2.9 :

Dada uma matriz de ordem $p \times p$, simétrica e positiva definida B , e um escalar $b > 0$,

$$\frac{1}{|\Sigma|^b} e^{-\text{tr}(\Sigma^{-1}B)/2} \leq \frac{1}{|B|^b} (2b)^{pb} e^{-bp}$$

para toda matriz positiva definida $\Sigma_{(p \times p)}$, com a igualdade verificando-se apenas para

$$\Sigma = (1/2b)B.$$

Prova:

Seja $B^{1/2}$ a matriz raiz quadrada de B , então $B^{1/2} \cdot B^{1/2} = B$, $B^{1/2} \cdot B^{-1/2} = I$ e $B^{-1/2} \cdot B^{1/2} = B^{-1}$. Assim, $\text{tr}(\Sigma^{-1}B) = \text{tr}[(\Sigma^{-1}B^{1/2})B^{1/2}] = \text{tr}[B^{1/2}(\Sigma^{-1}B^{1/2})]$. Seja η um autovalor de $B^{1/2}\Sigma^{-1}B^{1/2}$. Esta matriz é positiva definida pois $y'B^{1/2}\Sigma^{-1}B^{1/2}y = (B^{1/2}y)'\Sigma^{-1}(B^{1/2}y) > 0$ se $B^{1/2}y \neq 0$ ou da mesma forma se $y \neq 0$. Assim os autovalores η_i de $B^{1/2}\Sigma^{-1}B^{1/2}$ são positivos pelo resultado 2.6. O resultado 2.12 leva a :

$$\text{tr}(\Sigma^{-1}B) = \text{tr}(B^{1/2}\Sigma^{-1}B^{1/2}) = \sum_{i=1}^p \eta_i$$

e $|B^{1/2}\Sigma^{-1}B^{1/2}| = \prod_{i=1}^p \eta_i$ pelo resultado 2.8. Das propriedades de determinantes pode-se escrever

$$|B^{1/2}\Sigma^{-1}B^{1/2}| = |B^{1/2}| |\Sigma^{-1}| |B^{1/2}| = |\Sigma^{-1}| |B^{1/2}| |B^{1/2}| = |\Sigma^{-1}| |B| = \frac{1}{|\Sigma|} |B|$$

ou

$$\frac{1}{|\Sigma|} = \frac{|B^{1/2}\Sigma^{-1}B^{1/2}|}{|B|} = \frac{\prod_{i=1}^p \eta_i}{|B|}$$

combinando os resultados para o traço e determinante, tem-se

$$\frac{1}{|\Sigma|^b} e^{-\text{tr}[\Sigma^{-1}B]/2} = \frac{\left(\prod_{i=1}^p \eta_i\right) e^{\sum_{i=1}^p \eta_i/2}}{|B|^p} = \frac{1}{|B|^p} \prod_{i=1}^p \eta_i^b e^{-\eta_i/2}$$

Mas a função $\eta^b e^{-\eta/2}$ tem um máximo, em relação a η , de $(2b)^b e^{-b}$, ocorrendo em $\eta = 2b$.

A escolha $\eta_i = 2b$, para cada i , leva a

$$\frac{1}{|\Sigma|^b} e^{-\text{tr}[\Sigma^{-1}B]/2} \leq \frac{1}{|B|^p} (2b)^{pb} e^{-bp}$$

O limite superior é único e obtido quando $\Sigma = (1/2b)B$, desde que para esta escolha,

$$B^{1/2}\Sigma^{-1}B^{1/2} = B^{1/2} (2b) B^{-1} B^{1/2} = (2b) \underset{(p \times p)}{I}$$

e

$$\text{tr}(\Sigma^{-1}B) = \text{tr}(B^{1/2}\Sigma^{-1}B^{1/2}) = \text{tr}[(2b)I] = 2bp$$

além disso

$$\frac{1}{|\Sigma|} = \frac{|B^{1/2}\Sigma^{-1}B^{1/2}|}{|B|} = \frac{|(2b)I|}{|B|} = \frac{(2b)^p}{|B|}$$

a substituição direta para $\text{tr}[\Sigma^{-1}B]$ e $1/|\Sigma|^b$ leva ao limite previsto.

Resultado 2.10 :

Seja X_1, X_2, \dots, X_n uma amostra aleatória de uma população normal p variada, com média μ e covariância Σ . Então

$$\hat{\mu} = \bar{X} \quad \text{e} \quad \hat{\Sigma} = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})(X_j - \bar{X})' = \frac{(n-1)}{n} S$$

são, respectivamente, os estimadores de máxima verossimilhança de μ e Σ e seus valores

observados, \bar{x} e $(1/n) \sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})'$, são denominados estimativas de máxima

verossimilhança para μ e Σ , com $S = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})(X_j - \bar{X})'$.

Prova:

Com exceção do fator multiplicador $-1/2$, o expoente da função de verossimilhança é

$$\text{tr}[\Sigma^{-1} \left(\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})' \right)] + n \cdot (\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu)$$

Pelo resultado 2.5, Σ^{-1} é positiva definida, portanto, a distância $(\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu) > 0$, sendo nula apenas se $\mu = \bar{x}$. Assim, a verossimilhança é maximizada em relação a μ em $\hat{\mu} = \bar{x}$.

Resta maximizar

$$L(\mu, \Sigma) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\Sigma|^{n/2}} \cdot e^{-\text{tr} \left[\Sigma^{-1} \left(\sum_{j=1}^n (x_{\sim j} - \bar{x})(x_{\sim j} - \bar{x})' \right) \right] / 2}$$

sobre Σ . Pelo resultado 2.9, com $b = n/2$ e $B = \sum_{j=1}^n (x_{\sim j} - \bar{x})(x_{\sim j} - \bar{x})'$, o máximo ocorre

em $\hat{\Sigma} = (1/n) \sum_{j=1}^n (x_{\sim j} - \bar{x})(x_{\sim j} - \bar{x})'$ conforme foi declarado.

Os estimadores de máxima verossimilhança são quantidades aleatórias, que são obtidas pela troca das observações x_1, x_2, \dots, x_n nas expressões $\hat{\mu}$ e $\hat{\Sigma}$, pelos correspondentes vetores aleatórios X_1, X_2, \dots, X_n .

2.4.6 MÉTODOS DE ESTIMAÇÃO

O Modelo Fatorial Ortogonal procura representar adequadamente os dados com um pequeno número de fatores não observáveis. A matriz de covariância S é um estimador da matriz populacional Σ desconhecida. Se num determinado problema, os elementos fora da diagonal de S são pequenos, ou de forma equivalente, os elementos fora da diagonal da matriz R de correlação são essencialmente nulos, então as variáveis não são correlacionadas, e, portanto, uma Análise Fatorial não será útil ao problema. Nessas circunstâncias, os fatores

específicos realizam uma importante função na Análise Fatorial, já que a proposta principal é determinar poucos, mas importantes fatores comuns.

Se Σ desvia significativamente de uma matriz diagonal, então o Modelo Fatorial pode ser utilizado, e o problema inicial é uma estimação dos carregamentos fatoriais ℓ_{ij} e variâncias específicas ψ_i . Os métodos mais populares de estimação de parâmetros são: o Método de Componentes Principais e o Método de Máxima Verossimilhança. As soluções de qualquer um deles pode ser rotacionada para a obtenção de uma estrutura mais simples, na qual a interpretação dos fatores seja bem mais visível e simplificada. É sempre prudente numa análise de dados utilizar ambos os métodos de estimação. Dependendo do problema, um deles pode se adequar melhor, e as soluções podem ser bem mais consistentes.

2.4.6.1 Método de Componentes Principais

A matriz de covariância Σ pode ser fatorada em função de seus autovalores e autovetores através da decomposição espectral. Sendo assim, seja uma matriz de covariância Σ cujos pares de autovalor-autovetor têm a forma $(\lambda_i, e_{\sim i})$ com $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Então, a decomposição espectral da matriz de covariância pode ser escrita:

$$\Sigma = \lambda_1 e_{\sim 1} e_{\sim 1}' + \lambda_2 e_{\sim 2} e_{\sim 2}' + \dots + \lambda_p e_{\sim p} e_{\sim p}'$$

ou em uma forma matricial fatorada

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} \cdot e_{-1} & & & \\ & \sqrt{\lambda_2} \cdot e_{-2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_p} \cdot e_{-p} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{\lambda_1} \cdot e_{-1}' \\ \dots \\ \sqrt{\lambda_2} \cdot e_{-2}' \\ \dots \\ \sqrt{\lambda_p} \cdot e_{-p}' \end{bmatrix} \quad (\text{I})$$

A equação (I) descreve o comportamento de um modelo em que $m = p$, ou seja, o número de fatores é igual ao número de variáveis. Neste caso, comparativamente em relação à estrutura do Modelo Fatorial Ortogonal, a matriz de variâncias específicas Ψ passa a ser a matriz nula, uma vez que os fatores comuns explicam integralmente a porção da variância das variáveis aleatórias observáveis X . Portanto, a estimação seria perfeita, e não haveria nenhuma parcela de aleatoriedade atribuída aos fatores comuns. Então a matriz de covariância ficaria:

$$\underset{(pxp)}{\Sigma} = \underset{(pxp)}{L} \cdot \underset{(pxp)}{L}' + \underset{(pxp)}{\Psi} = \underset{(pxp)}{L} \cdot \underset{(pxp)}{L}' + \underset{(pxp)}{0} \Rightarrow \underset{(pxp)}{\Sigma} = \underset{(pxp)}{L} \cdot \underset{(pxp)}{L}'$$

e os fatores específicos ε não seriam considerados, pois apresentariam pouca importância no modelo. Caso se incluísse na fatoração de Σ , também os fatores específicos, eles seriam os elementos diagonais da matriz $\Sigma - LL'$, onde LL' é a decomposição espectral de Σ . As relações a seguir mostram isso.

$$\Sigma = LL' + \Psi$$

ou

$$\Psi = \Sigma - LL'$$

$$\begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_p \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix} - \begin{bmatrix} \sqrt{\lambda_1} \cdot e_{-1} & & & \\ & \sqrt{\lambda_2} \cdot e_{-2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_p} \cdot e_{-p} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{\lambda_1} \cdot e_{-1}' \\ \dots \\ \sqrt{\lambda_2} \cdot e_{-2}' \\ \dots \\ \sqrt{\lambda_p} \cdot e_{-p}' \end{bmatrix} \quad (\text{II})$$

onde $\psi_i = \sigma_{ii} - \sum_{j=1}^p \ell_{ij}^2$ para $i = 1, 2, \dots, p$.

Na aplicação do modelo descrito anteriormente, para que as observações $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ estejam centradas no mesmo ponto, costuma-se subtrair de cada observação a média amostral \bar{x} . Além disso, nos casos em que as unidades das variáveis são medidas em escalas com uma larga diferença, usualmente trabalha-se com as variáveis padronizadas, ou seja, após subtrair a média amostral, divide-se pelo desvio padrão de cada variável. A seguir tem-se a forma das variáveis padronizadas $z_{\tilde{j}}$.

$$z_{\tilde{j}} = \begin{bmatrix} \frac{(x_{1j} - \bar{x}_1)}{\sqrt{s_{11}}} \\ \frac{(x_{2j} - \bar{x}_2)}{\sqrt{s_{22}}} \\ \vdots \\ \frac{(x_{pj} - \bar{x}_p)}{\sqrt{s_{pp}}} \end{bmatrix}, j = 1, 2, \dots, n$$

Um problema que tenha uma variável com uma grande variância pode apresentar uma interpretação incorreta dos resultados. Essa variável pode influenciar indevidamente o modelo, distorcendo os resultados obtidos. Para evitar situações como essa, usualmente as variáveis são padronizadas.

Para variáveis padronizadas, a matriz S , que estima a matriz de covariância Σ , é trocada pela matriz de correlação amostral R , que estima a matriz de correlação populacional ρ .

A Análise de Componentes Principais de uma matriz de covariância amostral S é realizada em função de seus pares de autovalor-autovetor $(\hat{\lambda}_1, \hat{e}_{\sim 1}), (\hat{\lambda}_2, \hat{e}_{\sim 2}), \dots, (\hat{\lambda}_p, \hat{e}_{\sim p})$, onde $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p$. Sendo m o número de fatores comuns, a matriz dos carregamentos fatoriais estimados $\{\hat{\ell}_{ij}\}$ é dada por

$$\hat{L} = \left[\sqrt{\hat{\lambda}_1} \cdot \hat{e}_{\sim 1} \quad \sqrt{\hat{\lambda}_2} \cdot \hat{e}_{\sim 2} \quad \dots \quad \sqrt{\hat{\lambda}_m} \cdot \hat{e}_{\sim m} \right]$$

As variâncias específicas estimadas são produzidas pelos elementos diagonais da matriz $S - \hat{L}\hat{L}'$, então

$$\hat{\Psi} = \begin{bmatrix} \hat{\psi}_1 & 0 & \dots & 0 \\ 0 & \hat{\psi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\psi}_p \end{bmatrix} \text{ com } \hat{\psi}_i = s_{ii} - \sum_{j=1}^m \hat{\ell}_{ij}^2$$

As comunicações são estimadas através de

$$\hat{h}_i^2 = \hat{\ell}_{i1}^2 + \hat{\ell}_{i2}^2 + \dots + \hat{\ell}_{im}^2$$

A análise de componentes principais da matriz de correlação amostral é obtido substituindo R no lugar de S .

No estudo do Método de Coeficientes Principais, os primeiros coeficientes ponderados das componentes principais amostrais são os carregamentos fatoriais de (I). Por esse fato, é que se denomina o modelo descrito como Modelo de Componentes Principais.

2.4.6.2 Verossimilhança da Normal Multivariada

Em muitos casos, para problemas que envolvem populações normalmente distribuídas, pode-se usar o Método da Máxima Verossimilhança para estimar os carregamentos.

Sejam $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$, vetores com p componentes representando amostras aleatórias de uma população multivariada normal com média $\tilde{\mu}$ e matriz de covariância $\tilde{\Sigma}$, se esses vetores são independentes e cada um tem distribuição normal $N_p(\tilde{\mu}, \tilde{\Sigma})$, então a função de densidade conjunta de todas as observações é o produto das densidades normais, ou seja

$$p(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = p(\tilde{X}_1) \cdot p(\tilde{X}_2) \dots p(\tilde{X}_n)$$

$$p(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = \prod_{j=1}^n p(\tilde{X}_j)$$

ou

$$p(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = \prod_{j=1}^n \left\{ \frac{1}{(2\pi)^{p/2} |\tilde{\Sigma}|^{1/2}} \cdot e^{-\frac{1}{2}(\tilde{x}_j - \tilde{\mu})^T \tilde{\Sigma}^{-1} (\tilde{x}_j - \tilde{\mu})} \right\}$$

desenvolvendo o produtório e agrupando, tem-se:

$$p(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\tilde{\Sigma}|^{n/2}} \cdot e^{-\frac{1}{2} \sum_{j=1}^n (\tilde{x}_j - \tilde{\mu})^T \tilde{\Sigma}^{-1} (\tilde{x}_j - \tilde{\mu})} \quad (I)$$

Quando os valores numéricos das observações são disponíveis, podem ser substituídos no lugar de \tilde{x}_j na equação anterior. A expressão resultante que é dada como uma função de $\tilde{\mu}$ e $\tilde{\Sigma}$ para o conjunto de observações fixadas $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$, é denominada função de verossimilhança.

Muitos procedimentos estatísticos empregam estimativas de parâmetros populacionais para melhor explicar os dados observados. Uma técnica muito eficiente, quando aplicada a populações normais, consiste em selecionar um valor para o parâmetro que maximiza a densidade conjunta avaliada das observações. Esta técnica é denominada Estimação de Máxima Verossimilhança .

Denotando a função de verossimilhança por $L(\underline{\mu}, \underline{\Sigma})$, a equação (I) seria assim escrita:

$$L(\underline{\mu}, \underline{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underline{\Sigma}|^{n/2}} \cdot e^{-\sum_{j=1}^n (\underline{x}_j - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{x}_j - \underline{\mu}) / 2} \quad (\text{II})$$

Em muitos problemas, é conveniente expressar a função de verossimilhança de outra forma. O resultado 2.15 mostra que desmembrando o 2° membro e utilizando propriedades do traço de matrizes quadradas, (II) pode ser colocada na seguinte maneira:

$$L(\underline{\mu}, \underline{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underline{\Sigma}|^{n/2}} \cdot e^{-\frac{1}{2} \text{tr} \left[\underline{\Sigma}^{-1} \left(\sum_{j=1}^n (\underline{x}_j - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})' \right) \right] + \frac{n}{2} \cdot (\underline{\bar{x}} - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{\bar{x}} - \underline{\mu})}$$

Resultado 2.11:

Seja A uma matriz simétrica de ordem $k \times k$ e \underline{x} um vetor de dimensão k . Então

$$\underline{x}' A \underline{x} = \text{tr}(\underline{x}' A \underline{x}) = \text{tr}(A \underline{x} \underline{x}')$$

Prova:

Pode-se notar que $\underline{x}' A \underline{x}$ é um escalar, logo $\underline{x}' A \underline{x} = \text{tr}(\underline{x}' A \underline{x})$. Assim a primeira igualdade é imediata.

Da propriedade do traço de uma matriz quadrada, tem-se que $\text{tr}(BC) = \text{tr}(CB)$.

Supondo B e C matrizes de ordem mxk e kxm , respectivamente, o i -ésimo elemento diagonal

de BC é $\sum_{j=1}^k b_{ij}c_{ji}$. Portanto, o traço de BC é dado por $\text{tr}(BC) = \sum_{i=1}^m \left(\sum_{j=1}^k b_{ij}c_{ji} \right)$. Similarmente o

j -ésimo elemento da diagonal de CB é $\sum_{i=1}^m c_{ji}b_{ij}$, então

$$\text{tr}(CB) = \sum_{j=1}^k \left(\sum_{i=1}^m c_{ji}b_{ij} \right) = \sum_{i=1}^m \left(\sum_{j=1}^k b_{ij}c_{ji} \right) = \text{tr}(BC).$$

Tomando $m = 1$ para a matriz B e trocando a matriz C por Ax , facilmente seria provado que

$$\text{tr}(x'(Ax)) = \text{tr}((Ax)x').$$

Resultado 2.12 :

$$\text{tr}(A) = \sum_{i=1}^k \lambda_i, \text{ onde } \lambda_i \text{ são os autovalores de } A \text{ de ordem } k.$$

Prova:

Usando a decomposição espectral, a matriz A pode ser escrita $A = P\Lambda P$, onde P é uma matriz ortogonal e Λ é uma matriz diagonal, cujos elementos são os autovalores da matriz A . Assim, $\text{tr}(A) = \text{tr}(P'\Lambda P) = \text{tr}(\Lambda P P') = \text{tr}(\Lambda \cdot I) = \text{tr}(\Lambda) = \lambda_1 + \lambda_2 + \dots + \lambda_k$.

Resultado 2.13:

A função de verossimilhança para populações normais independentes e aleatoriamente distribuídas é dada por

$$L(\underline{\mu}, \underline{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underline{\Sigma}|^{n/2}} \cdot e^{-\frac{1}{2} \text{tr} \left[\underline{\Sigma}^{-1} \left(\sum_{j=1}^n (\underline{x}_j - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})' \right) \right] + \frac{n}{2} (\underline{\bar{x}} - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{\bar{x}} - \underline{\mu})}$$

Prova:

A função de verossimilhança para populações normais independentes aleatoriamente distribuídas é dada por

$$L(\underline{\mu}, \underline{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underline{\Sigma}|^{n/2}} \cdot e^{-\sum_{j=1}^n (\underline{x}_j - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{x}_j - \underline{\mu}) / 2}$$

O expoente da densidade conjunta pode ser simplificado. Levando-se em conta que $(\underline{x}_j - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{x}_j - \underline{\mu})$ é um escalar, de acordo com o resultado 2.11, pode-se escrever:

$$\begin{aligned} \sum_{j=1}^n (\underline{x}_j - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{x}_j - \underline{\mu}) &= \sum_{j=1}^n \text{tr}[(\underline{x}_j - \underline{\mu})' \underline{\Sigma}^{-1} (\underline{x}_j - \underline{\mu})] = \sum_{j=1}^n \text{tr}[\underline{\Sigma}^{-1} (\underline{x}_j - \underline{\mu})(\underline{x}_j - \underline{\mu})'] = \\ &= \text{tr}[\underline{\Sigma}^{-1} \left(\sum_{j=1}^n (\underline{x}_j - \underline{\mu})(\underline{x}_j - \underline{\mu})' \right)] \end{aligned}$$

Adicionando e subtraindo $\underline{\bar{x}} = (1/n) \sum_{j=1}^n \underline{x}_j$ a cada termo $(\underline{x}_j - \underline{\mu})$ de

$\sum_{j=1}^n (\underline{x}_j - \underline{\mu})(\underline{x}_j - \underline{\mu})'$ obtém-se:

$$\begin{aligned} \text{tr}[\underline{\Sigma}^{-1} \left(\sum_{j=1}^n (\underline{x}_j - \underline{\bar{x}} + \underline{\bar{x}} - \underline{\mu})(\underline{x}_j - \underline{\bar{x}} + \underline{\bar{x}} - \underline{\mu})' \right)] &= \text{tr}[\underline{\Sigma}^{-1} \left(\sum_{j=1}^n (\underline{x}_j - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})' + \sum_{j=1}^n (\underline{\bar{x}} - \underline{\mu})(\underline{\bar{x}} - \underline{\mu})' \right)] = \\ &= \text{tr}[\underline{\Sigma}^{-1} \left(\sum_{j=1}^n (\underline{x}_j - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})' + n(\underline{\bar{x}} - \underline{\mu})(\underline{\bar{x}} - \underline{\mu})' \right)] = \end{aligned}$$

$$\begin{aligned}
&= \text{tr}[\Sigma^{-1} \left(\sum_{j=1}^n (\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})' \right)] + n \cdot \text{tr}[\Sigma^{-1} (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})'] = \\
&= \text{tr}[\Sigma^{-1} \left(\sum_{j=1}^n (\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})' \right)] + n \cdot (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})' \Sigma^{-1} (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})]
\end{aligned}$$

Portanto, a função de verossimilhança em vez de

$$L(\underset{\sim}{\mu}, \underset{\sim}{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underset{\sim}{\Sigma}|^{n/2}} \cdot e^{-\sum_{j=1}^n (\underset{\sim}{x}_j - \underset{\sim}{\mu})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{x}_j - \underset{\sim}{\mu}) / 2}$$

pode ser escrita

$$L(\underset{\sim}{\mu}, \underset{\sim}{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underset{\sim}{\Sigma}|^{n/2}} \cdot e^{-\frac{1}{2} \cdot \text{tr} \left[\Sigma^{-1} \left(\sum_{j=1}^n (\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})' \right) \right] + \frac{n}{2} \cdot (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})' \Sigma^{-1} (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})}$$

2.4.6.3 Método da Máxima Verossimilhança

Na Análise Fatorial, se os fatores comuns $\underset{\sim}{F}$ e fatores específicos $\underset{\sim}{\varepsilon}$ são considerados normalmente distribuídos, então as estimativas dos carregamentos fatoriais e variâncias específicas podem ser obtidas. Quando $\underset{\sim}{F}_j$ e $\underset{\sim}{\varepsilon}_j$ são conjuntamente normais, as observações

$\underset{\sim}{X}_j - \underset{\sim}{\mu} = L\underset{\sim}{F}_j + \underset{\sim}{\varepsilon}_j$ são também normais. Assim sendo, a verossimilhança de populações normalmente distribuídas, através do resultado 2.13, tem-se:

$$L(\underset{\sim}{\mu}, \underset{\sim}{\Sigma}) = \frac{1}{(2\pi)^{np/2}} \cdot \frac{1}{|\underset{\sim}{\Sigma}|^{n/2}} \cdot e^{-\frac{1}{2} \cdot \text{tr} \left[\Sigma^{-1} \left(\sum_{j=1}^n (\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})' \right) \right] + \frac{n}{2} \cdot (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})' \Sigma^{-1} (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})}$$

Desmembrando as somas nos expoentes em produtos de potências na mesma base, tem-se:

$$L(\mu, \Sigma) = (2\pi)^{-\frac{(n-1)p}{2}} \cdot |\Sigma|^{-\frac{(n-1)}{2}} \cdot e^{-\left(\frac{1}{2}\right) \left[\sum_{j=1}^n (x_j - \bar{x})(x_j - \bar{x})' \right]} \cdot (2\pi)^{-\frac{p}{2}} \cdot |\Sigma|^{-\frac{1}{2}} \cdot e^{-\left(\frac{n}{2}\right) (\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu)}$$

Já que $\Sigma = LL' + \Psi$, a função de verossimilhança depende de L e Ψ . Pela multiplicidade das escolhas de L , possíveis por transformações ortogonais, é conveniente impor uma condição de singularidade computacional. A condição de singularidade é que

$$L' \Psi^{-1} L = \Delta$$

seja uma matriz diagonal.

Pela propriedade da invariância de estimadores de máxima verossimilhança, funções de L e Ψ são estimadas pelas mesmas funções de \hat{L} e $\hat{\Psi}$. Desta forma, a máxima verossimilhança estima que as comunalidades são

$$\hat{h}_i^2 = \hat{\ell}_{i1}^2 + \hat{\ell}_{i2}^2 + \dots + \hat{\ell}_{im}^2 \quad \text{para } i = 1, 2, \dots, p$$

assim

$$\left(\begin{array}{l} \text{Porção do total da variância} \\ \text{amostral atribuída ao } j\text{-ésimo fator} \end{array} \right) = \frac{\hat{\ell}_{1j}^2 + \hat{\ell}_{2j}^2 + \dots + \hat{\ell}_{pj}^2}{s_{11} + s_{22} + \dots + s_{pp}}$$

Da mesma forma, como foi usada a matriz de covariância para estimar a matriz de carregamentos e a matriz de variâncias específicas, poderia ter sido usada a matriz de correlação. Para tanto, bastava que as variáveis fossem padronizadas.

As variáveis padronizadas Z tem a forma $Z = V^{-1/2}(X - \mu)$. Da relação entre as matrizes de correlação e covariância pode-se escrever:

$$\rho = V^{-1/2} \Sigma V^{-1/2}$$

Substituindo $\Sigma = LL' + \Psi$ na equação anterior :

$$\rho = (V^{1/2}L)(V^{1/2}L)' + V^{1/2}\Psi V^{1/2}$$

Assim, ρ foi fatorada de maneira análoga ao Modelo Fatorial Ortogonal da matriz de covariância. Com matriz de carregamentos $L_z = V^{1/2}L$ e matriz de variância específica $\Psi_z = V^{1/2}\Psi V^{1/2}$. Pela propriedade da invariância dos estimadores de máxima verossimilhança, o estimador de máxima verossimilhança de ρ é

$$\hat{\rho} = \left(\hat{V}^{-1/2} \hat{L} \right) \left(\hat{V}^{-1/2} \hat{L} \right)' + \hat{V}^{-1/2} \hat{\Psi} \hat{V}^{-1/2}$$

ou

$$\hat{\rho} = \hat{L}_z \hat{L}_z' + \hat{\Psi}_z$$

Como consequência desta fatoração, sempre que a análise de máxima verossimilhança é aplicada à matriz de correlação, a importância dos fatores é avaliada com base nos elementos da matriz de carregamentos padronizada \hat{L}_z . Desta forma, a porção da variância atribuída a um fator é dada por:

$$\left(\begin{array}{l} \text{Porção do total da variância amostral} \\ \text{padronizada atribuída ao } j\text{-ésimo fator} \end{array} \right) = \frac{\hat{\ell}_{1j}^2 + \hat{\ell}_{2j}^2 + \dots + \hat{\ell}_{pj}^2}{p}$$

onde $\hat{\ell}_{ij}$ representam os carregamentos de \hat{L}_z .

Ordinariamente as observações são padronizadas e a matriz de correlação amostral é fatorada e analisada. A matriz de correlação amostral R , é trocada por $[(n-1)/n]S$ na função de verossimilhança (II), e as estimativas de máxima verossimilhança \hat{L}_z e $\hat{\Psi}_z$ são obtidas usando um computador. Embora a verossimilhança (II) seja apropriada para S , não R , surpreendentemente, esta prática é equivalente a obter as estimativas de máxima verossimilhança \hat{L} e $\hat{\Psi}$ baseadas na matriz de covariância amostral S , fixando $\hat{L}_z = V^{-1/2} \hat{L}$ e

$\hat{\Psi}_z = \hat{V}^{-1/2} \hat{\Psi} \hat{V}^{-1/2}$. Neste caso, $\hat{V}^{-1/2}$ é a matriz diagonal, cuja diagonal principal é formada pelo inverso dos desvios padrão amostral, computado com o divisor \sqrt{n} .

Por outro lado, dados os carregamentos estimados \hat{L}_z e as variâncias específicas $\hat{\Psi}_z$ obtidas da matriz de correlação amostral R , as estimativas resultantes de máxima verossimilhança para a análise fatorial da matriz de covariância $[(n-1)/n]S$ são $\hat{L} = \hat{V}^{-1/2} \hat{L}_z$ e $\hat{\Psi} = \hat{V}^{-1/2} \hat{\Psi}_z \hat{V}^{-1/2}$, ou $\hat{\ell}_{ij} = \hat{z} \hat{\ell}_{ij} \sqrt{\hat{\sigma}_{ii}}$ e $\hat{\psi}_i = \hat{z} \hat{\psi}_i \hat{\sigma}_{ii}$ onde $\hat{\sigma}_{ii}$ é a variância amostral calculada com divisor n . A distinção entre divisores pode ser ignorada na solução de componentes principais.

A equivalência entre a fatoração de S e R vem aparentemente sendo confundida em muitas publicações.

2.4.6.4 Método do Fator Principal

Além dos métodos de estimação baseados em componentes principais e máxima verossimilhança, pode-se também estimar os carregamentos através do Método do Fator Principal. Este método, que é considerado uma modificação do Método de Componentes Principais, procura melhorar o estimador de L obtido das componentes principais, pelo uso de estimadores iniciais dos elementos diagonais de $S - \hat{\Psi}_i$ ou $R - \hat{\ell}_{ij}$ pela fatoração dessas matrizes em $\tilde{L}\tilde{L}'$:

$$S - \hat{\Psi} \cong \tilde{L}\tilde{L}'$$

$$R - \tilde{\Psi} \cong \tilde{L}\tilde{L}'$$

onde \tilde{L} tem ordem $p \times m$. Assim, se o Modelo Fatorial $\rho = LL' + \Psi$ está corretamente especificado, os m fatores comuns podem ser contabilizados pelos elementos fora da diagonal de ρ tão bem quanto os elementos da própria diagonal.

Será descrito a seguir o procedimento em termos da análise fatorial de R , mesmo que também se possa aplicar apropriadamente a S .

Os elementos diagonais de $R - \tilde{\Psi}$ são $\tilde{h}_i^2 = 1 - \tilde{\Psi}_i$ $i = 1, 2, \dots, p$. Se a contribuição do fator específico ψ_i é removida da diagonal ou, equivalentemente, troca-se 1 por \tilde{h}_i^2 , a matriz resultante é $\rho - \Psi = LL'$. Para o procedimento, primeiramente supõe-se que estimativas iniciais $\tilde{\psi}_i$ das variâncias específicas estejam disponíveis. Trocando o i -ésimo elemento diagonal de R por $\tilde{h}_i^2 = 1 - \tilde{\Psi}_i$, obtém-se R_r , que é uma matriz de correlação amostral “reduzida”:

$$R_r = \begin{bmatrix} \tilde{h}_1^{*2} & r_{12} & \cdots & r_{1p} \\ r_{12} & \tilde{h}_2^{*2} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1p} & r_{2p} & \cdots & \tilde{h}_p^{*2} \end{bmatrix}$$

Desta maneira, sem contar a variação amostral, todos os elementos da matriz de correlação amostral reduzida podem ser contabilizados pelos m fatores comuns. Em particular, R_r é fatorada da seguinte forma:

$$R_r = L_r^* L_r^{*'}$$

onde $L_r^* = \{\ell_{ij}^*\}$ são os carregamentos estimados.

O Método do Fator Principal da Análise Fatorial emprega as estimativas

$$L_r^* = \left[\sqrt{\hat{\lambda}_{r1}^*} \hat{e}_{\sim r1}^* \quad \sqrt{\hat{\lambda}_{r2}^*} \hat{e}_{\sim r2}^* \quad \dots \quad \sqrt{\hat{\lambda}_{rm}^*} \hat{e}_{\sim rm}^* \right]$$

onde $(\hat{\lambda}_{ri}^*, \hat{e}_{\sim ri}^*)$, $i = 1, 2, \dots, m$ são os maiores pares de autovalor-autovetor determinados de R_r . É importante ressaltar que são considerados apenas os pares de autovalor-autovetor, cujos autovalores explicam a maior parte da variância possível, dependendo do critério adotado, conforme item 2.4.5. Em consequência as comunalidades seriam reestimadas por

$$h_i^{*2} = \sum_{j=1}^m \ell_{ij}^{*2}$$

A solução do Fator Principal pode ser obtido iterativamente, com as comunalidades estimadas pela equação anterior, que se tornam as estimativas iniciais para o próximo passo.

Os autovalores estimados $\hat{\lambda}_1^*, \hat{\lambda}_2^*, \dots, \hat{\lambda}_p^*$ auxiliam na determinação do número de fatores comuns na análise. Uma complicação adicional acontece pelo fato de que os autovalores podem ser negativos, pois foram consideradas as comunalidades estimadas iniciais. O ideal nesses casos é considerar o número de fatores igual ao posto da matriz populacional reduzida. De qualquer forma, o posto não é sempre bem determinado de R_r e algumas considerações adicionais são necessárias.

Mesmo que existam muitas escolhas para as estimativas das variâncias específicas iniciais, a escolha mais usada, quando se considera a matriz de correlação, é $\psi_i^* = \frac{1}{r_{ii}}$, no qual r^{ii} é o i -ésimo elemento diagonal de R^{-1} . As comunalidades iniciais seriam dadas então por

$$h_i^{*2} = 1 - \psi_i^* = 1 - \frac{1}{r^{ii}}$$

Essas estimativas representam o coeficiente de correlação múltipla quadrado entre X_i e as outras $p-1$ variáveis. A relação do coeficiente de correlação múltipla significa que h_i^{*2} pode ser calculado sempre que R não tenha posto cheio.

Mesmo que o Método de Componentes Principais para R possa ser considerado como o Método do Fator Principal com estimativas unitárias das comunalidades iniciais, ou nulas das variâncias específicas, os dois são filosoficamente e geometricamente diferentes. Na prática, entretanto, freqüentemente produzem carregamentos fatoriais similares, se o número de variáveis é grande e o número de fatores comuns é pequeno.

De uma maneira geral, recomenda-se, com maior eficiência, o uso do Método de Componentes Principais ou Método da Máxima Verossimilhança.

2.4.6.5 Outros Métodos de Estimação

Na Análise Fatorial, muitos métodos podem ser descritos para a estimação de carregamentos e comunalidades. Dentre eles, além dos já descritos Método de Componentes Principais e Método da Máxima Verossimilhança, destacam-se também o Método Centróide, a Análise Fatorial Canônica, o Método de Mínimos Resíduos, a Análise de Imagem e a Fatoração Alfa. A seguir, algumas características desses métodos.

O Método Centróide é computacionalmente uma simplificação aproximada do Método do Fator Principal. Antes da era da computação eletrônica era muito utilizado. Atualmente, reserva-se a apenas um interesse histórico.

A Análise Fatorial Canônica procura fatores que apresentem um maior grau de relacionamento com as variáveis. Além disso, o método é invariante para diferentes escalas.

O Método de Mínimos Resíduos ou método de “mineiros” procura minimizar a soma dos quadrados das diferenças entre as correlações observadas e as esperadas do modelo.

A Análise de Imagem diferencia a imagem da variável da anti-imagem. A imagem é o valor esperado da variável, e a anti-imagem é o erro ou resíduo na regressão. Este método é invariante para diferentes escalas.

A Fatoração Alfa procura fatores que tenham correlação máxima com os correspondentes fatores, no universo das variáveis. Como outros, também é invariante para diferentes escalas.

Em 1990, KAISER e DERFLINGER (citados por RENCHER, 1998, p. 385), no trabalho intitulado “*Some contrasts between maximum likelihood factor analysis and alpha factor analysis*”, compararam a Análise Fatorial de Máxima Verossimilhança e a Análise Fatorial Alfa em termos das suposições das distribuições, complexidade computacional e outras propriedades. No trabalho, a Análise Fatorial Alfa mostrou-se superior.

Em trabalhos datados de 1982 e 1983, intitulados “*EM algorithms for ML factor analysis*” e “*More on EM for ML factor analysis*”, RUBIN e THAYER (citados por RENCHER, 1998, p. 385) discutiram o uso do algoritmo EM (Esperança – Maximização) na estimação de máxima verossimilhança dos carregamentos fatoriais. Basicamente, o algoritmo EM apresenta dois passos importantes. No passo da esperança, o algoritmo estima um valor que falta nos dados por médias de esperanças condicionadas, dados os valores observados do vetor e os parâmetros estimados. No passo da maximização, o algoritmo usa os valores estimados que faltam e calcula as estimativas dos parâmetros por médias da máxima verossimilhança. O processo iterativo continua até que os parâmetros dos valores estimados convirjam.

Atualmente, estudos têm sido realizados, com a aplicação em diversas áreas, na busca de novos métodos e com o intuito de melhorar a estimação de carregamentos fatoriais e comunalidades. Mesmo assim, de uma forma geral, os mais usados ainda são o Método de Componentes Principais e o Método da Máxima Verossimilhança.

2.4.7 ROTAÇÃO DE FATORES

Os carregamentos do Modelo Fatorial Ortogonal do item 2.4.3 podem ser multiplicados por uma matriz ortogonal T sem prejudicar a reprodução da matriz de covariância em $\Sigma = LL' + \Psi$. Da Álgebra Matricial, sabe-se que uma transformação ortogonal corresponde a uma rotação de eixos. Por esta razão, essa transformação é denominada rotação fatorial.

Quando o número de fatores da análise é relativamente pequeno, mais precisamente 2, a separação dos grupos das variáveis em função dos carregamentos fatoriais é relativamente simples e pode até ser estabelecida por um método gráfico que torne a matriz de carregamentos uma estrutura simples. Nesse caso, os fatores não correlacionados seriam os eixos, e cada um dos p pontos da forma $\left(\hat{\ell}_{i1}, \hat{\ell}_{i2} \right)$ representaria uma variável. Os novos eixos fatoriais seriam então rotacionados de um ângulo θ em torno da origem para que os novos carregamentos rotacionados $\hat{\ell}_{ij}^*$ sejam obtidos. Como uma transformação ortogonal é uma transformação linear, utilizando o resultado 2.17 e da trigonometria elementar, as imagens dos vetores-coluna unitários seriam dados por

$$T(e_{-1}) = (\cos\theta, -\text{sen}\theta)$$

e, no caso de uma rotação no sentido horário

$$T(e_{-2}) = \left(\cos\left(\theta - \frac{\pi}{2}\right), -\text{sen}\left(\theta - \frac{\pi}{2}\right) \right) = (\text{sen}\theta, \cos\theta)$$

ou

$$T(e_1) = (\cos\theta, \text{sen}\theta)$$

e, no caso de uma rotação no sentido horário

$$T(e_{-2}) = \left(\cos\left(\theta + \frac{\pi}{2}\right), \text{sen}\left(\theta + \frac{\pi}{2}\right) \right) = (-\text{sen}\theta, \cos\theta)$$

Assim sendo, o relacionamento entre os carregamentos originais e os novos carregamentos (rotacionados) para $m=2$ pode ser determinado através de:

$$\hat{L}_{(px2)}^* = \hat{L}_{(px2)} \cdot \hat{T}_{(2x2)}$$

Para uma rotação no sentido horário, a matriz ortogonal T correspondente a transformação linear seria

$$T = \begin{bmatrix} \cos\theta & \text{sen}\theta \\ -\text{sen}\theta & \cos\theta \end{bmatrix}$$

Para uma rotação no sentido anti-horário, a matriz ortogonal T correspondente a transformação linear seria

$$T = \begin{bmatrix} \cos\theta & -\text{sen}\theta \\ \text{sen}\theta & \cos\theta \end{bmatrix}$$

O ângulo θ é escolhido de forma a maximizar a variância. Conseguem-se isto deixando as variáveis com poucos carregamentos com coeficientes altos e os demais carregamentos com baixos coeficientes ou até mesmo nulos. A Figura 3 a seguir ilustra claramente o papel da uma rotação de fatores.

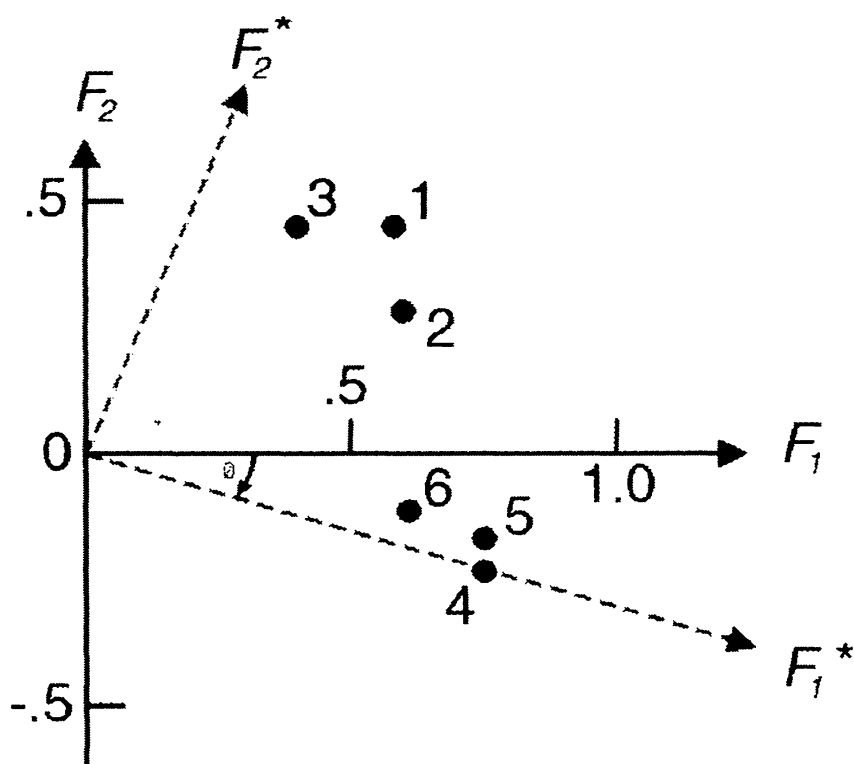


FIGURA 3 - ROTAÇÃO DE FATORES

As variáveis são localizadas pelos pontos 1, 2, 3, 4, 5 e 6. Os eixos representam dois fatores de um Modelo Fatorial Ortogonal. Antes da rotação, não há clareza em relação às variáveis quanto ao agrupamento nos fatores. Após a rotação, com os novos eixos F_1^* e F_2^* , as variáveis agrupadas em cada fator são definidas de forma cristalina. Desta forma, os carregamentos rotacionados seriam definidos com as variáveis 4, 5 e 6 agrupadas no 1º fator e as variáveis 1, 2 e 3 agrupadas no 2º fator.

Já para um número de fatores maior do que 2, nem sempre é fácil obter uma imediata interpretação. A estrutura simples que se visa com a rotação pode apresentar problemas, tais como muitos carregamentos fatoriais com grandes coeficientes para uma mesma variável, ou mesmo carregamentos com muitos coeficientes negativos, nos quais nenhuma variável se

destaca no fator. Mesmo assim, o método gráfico poderia ser usado. Neste caso, as variáveis deveriam ser localizadas em gráficos no IR^2 , onde os eixos seriam os fatores que seriam escolhidos dois a dois. A cada representação gráfica, os eixos seriam rotacionados e encontrariam-se, assim, melhores interpretações a cada par de fatores. O inconveniente do método gráfico aplicado sucessivamente é o imenso trabalho despendido pelas inúmeras possibilidades de escolhas dos fatores. Para evitar excessivo trabalho, foram desenvolvidos métodos computacionalmente intensivos para a finalidade única da rotação de fatores. Dentre os vários métodos existentes, atualmente o mais usado é o Método Varimax ou Método Varimax Normal. Este método será descrito com detalhes no presente estudo.

A matriz ortogonal satisfaz a condição $TT^T = I$. Assim sendo, pode-se inseri-la no modelo $\tilde{X} - \tilde{\mu} = L\tilde{F} + \tilde{\varepsilon}$. Levando assim a:

$$\tilde{X} - \tilde{\mu} = LTT^T\tilde{F} + \tilde{\varepsilon}$$

$$\tilde{X} - \tilde{\mu} = (LT)(T^T\tilde{F}) + \tilde{\varepsilon}$$

Considerando $L^* = LT$ e $\tilde{F}^* = T^T\tilde{F}$, pode-se escrever:

$$\tilde{X} - \tilde{\mu} = L^*\tilde{F}^* + \tilde{\varepsilon}$$

Os novos fatores \tilde{F}^* satisfazem as suposições iniciais do modelo:

$$(i) \quad E(\tilde{F}^*) = 0$$

$$(ii) \quad Cov(\tilde{F}^*) = I$$

$$(iii) \quad Cov(\tilde{F}^*, \tilde{\varepsilon}) = 0.$$

As comunalidades $h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2$, $i = 1, 2, \dots, p$ ficam inalteradas com a transformação $L^* = LT$. Denotando a i -ésima linha de L por ℓ_i' , a comunalidade é dada por $h_i^2 = \ell_i' \ell_i$. A i -ésima linha de $L^* = LT$ é

$$\ell_i^{*'} = \ell_i' T$$

transpondo-se ambos os membros

$$\ell_i^{*'} = (\ell_i' T)' = T' \ell_i$$

Em termos de ℓ_i^{*} , a comunalidade é

$$h_i^{2*} = \ell_i^{*'} \ell_i^{*} = \ell_i' T T' \ell_i = \ell_i' I \ell_i = \ell_i' \ell_i = h_i^2$$

Assim, as comunalidades são as mesmas para os novos carregamentos. É interessante observar ainda que, $h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 = \ell_i' \ell_i$ é a distância da origem ao ponto de coordenadas $\ell_i' = (\ell_{i1}, \ell_{i2}, \dots, \ell_{im})$ no espaço m -dimensional dos carregamentos fatoriais. Como a multiplicação ortogonal é equivalente a uma rotação (ou reflexão) de eixos, $\ell_i' \ell_i$ e $\ell_i^{*'} \ell_i^{*}$ representam pontos diferentes. No entanto, ambos com a mesma distância até a origem. As variâncias específicas também são invariantes com a rotação. Desta forma, na Análise Fatorial, diferente de outras técnicas multivariadas, as variáveis observadas não são afetadas pela rotação dos coeficientes. Mudam apenas os eixos de referência.

Definição 2.14:

Se V e W são espaços vetoriais, então a função $T: V \rightarrow W$ é uma transformação linear, contanto que

$$(i) \quad T(\underline{u} + \underline{v}) = T(\underline{u}) + T(\underline{v})$$

$$(ii) \quad T(c\underset{\sim}{u}) = cT(\underset{\sim}{u})$$

Para todos os vetores de $\underset{\sim}{u}$ e $\underset{\sim}{v}$ de V e para todo escalar c .

Resultado 2.15:

A função $T : IR^n \rightarrow IR^m$ é uma transformação linear se e somente se ela é uma transformação matricial, e nesse caso a matriz A correspondente a T é dada por

$$A = [T(\underset{\sim}{e}_1) \quad T(\underset{\sim}{e}_2) \quad \dots \quad T(\underset{\sim}{e}_n)]$$

Em consequência o j -ésimo vetor coluna de A é a imagem $T(\underset{\sim}{e}_j)$ do j -ésimo vetor unitário $\underset{\sim}{e}_j$ da base canônica.

Prova:

Se $T : IR^n \rightarrow IR^m$ é uma transformação matricial definida por $T(\underset{\sim}{x}) = A\underset{\sim}{x}$, então para $\underset{\sim}{u}$ e $\underset{\sim}{v}$ vetores e c escalar, tem-se

$$T(\underset{\sim}{u} + \underset{\sim}{v}) = A(\underset{\sim}{u} + \underset{\sim}{v}) = A\underset{\sim}{u} + A\underset{\sim}{v} = T(\underset{\sim}{u}) + T(\underset{\sim}{v})$$

e

$$T(c\underset{\sim}{u}) = A(c\underset{\sim}{u}) = c(A\underset{\sim}{u}) = cT(\underset{\sim}{u})$$

Logo T é linear. Além disso, segue imediatamente da definição de multiplicação matricial que $T(\underset{\sim}{e}_j) = A\underset{\sim}{e}_j$ é o j -ésimo vetor-coluna de A .

Reciprocamente, suponha que $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ seja uma transformação linear. Seja a matriz A do tipo $m \times n$ definida por

$$A = [T(\underline{e}_1) \quad T(\underline{e}_2) \quad \dots \quad T(\underline{e}_n)]$$

Resta provar que $T(\underline{x}) = A \underline{x}$ para todo vetor $\underline{x} = (x_1, x_2, \dots, x_n)$ de \mathbb{R}^n e, dessa forma, que T é a transformação matricial correspondente a A . Usando a linearidade de T , pode-se escrever

$$T(\underline{x}) = T(x_1 \underline{e}_1 + x_2 \underline{e}_2 + \dots + x_n \underline{e}_n)$$

$$T(\underline{x}) = x_1 T(\underline{e}_1) + x_2 T(\underline{e}_2) + \dots + x_n T(\underline{e}_n)$$

$$T(\underline{x}) = \begin{bmatrix} T(\underline{e}_1) & T(\underline{e}_2) & \dots & T(\underline{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$T(\underline{x}) = A \underline{x}$$

Assim, completa-se a prova do resultado.

2.4.7.1 Critério Quartimax

Dentro do estudo da organização da habilidade mental, que deu origem à Análise Fatorial, o primeiro critério analítico para a determinação psicológica de fatores interpretáveis foi devido à CARROL (citado por KAYSER, 1958, p. 187) em um trabalho intitulado “*An analytical solution for approximating simple structure in factor analysis*”, de 1953. A

intenção era produzir uma explicação matemática para a estrutura simples de Thurstone. Esta estrutura simples baseava-se na idéia, através de algumas regras, de que a matriz de carregamentos fatoriais poderia ser representada por uma simples estrutura. Estrutura essa de fácil interpretação. Carrol sugeriu que para uma dada matriz fatorial,

$$f = \sum_{s < t} \sum_i \ell_{is}^2 \ell_{it}^2 \quad (I)$$

deve ser um mínimo, onde $i = 1, 2, \dots, p$ são variáveis, $s, t = 1, 2, \dots, m$ são fatores, e ℓ_{is} é o carregamento fatorial da i -ésima variável no s -ésimo fator. Ao que tudo indica, Carrol foi motivado a escrever (I) depois de uma inspeção profunda das cinco regras formais para estrutura simples de Thurstone.

As cinco regras do critério de Thurstone para uma estrutura simples foram elaboradas em 1947 e descritas em seu trabalho “*Multiple-factor analysis*”. A seguir são descritas as cinco regras originais:

- a) cada linha da matriz fatorial deve ter no mínimo um zero;
- b) se existem m fatores comuns, cada coluna da matriz fatorial deve ter no mínimo m zeros;
- c) para cada par de colunas da matriz fatorial deve haver muitas variáveis, cujos coeficientes sejam nulos em uma coluna mas não na outra;
- d) para cada par de colunas da matriz fatorial, uma grande parte das variáveis deve ter coeficientes quase nulos em ambas as colunas, quando existem quatro ou mais fatores;
- e) para cada par de colunas da matriz fatorial deve haver apenas um pequeno número de variáveis com coeficientes não nulos em ambas as colunas.

No trabalho original de Carrol, já citado anteriormente, ele produziu dois exemplos numéricos da aplicação de seu método. Sem a restrição de ortogonalidade, suas ilustrações

tiveram alguns resultados equivocados. Mesmo a aplicação de (I) parecendo trazer uma finalidade para a desejada estrutura simples, o critério teve grandes problemas em relação a sua aceitação, sendo também fortemente influenciado pela fatoração de testes complexos.

Na época de grandes desenvolvimentos, o critério de Carrol poderia não ter sido aceito, mas, pelo contrário, ainda procurou quebrar a rígida regra quantitativa e arbitrária da estrutura simples de Thurstone.

Quase que simultaneamente ao desenvolvimento de Carrol, NEUHAUS e WRIGLEY (citados por KAYSER, 1958, p. 188), no trabalho “*The quartimax method: an analytical approach to orthogonal simple structure*”, de 1954, SAUNDERS (op. cit., p. 188) em “*An analytic method for rotation to orthogonal simple structure*”, e FERGUSON (op. cit., p. 188), em “*The concept of parsimony in factor analysis*” de 1954, propuseram o que é usualmente denominado de Método Quartimax para uma estrutura simples ortogonal. Neuhaus e Wrigley sugeriram que a facilidade na interpretação da matriz fatorial, no estudo da estrutura simples, poderia ser encontrada quando a variância de todos os $p.m$ carregamentos quadrados da matriz fatorial fosse um máximo, ou seja

$$q_1 = \left[pm \sum_i \sum_s (\ell_{is}^2)^2 - \left(\sum_i \sum_s \ell_{is}^2 \right)^2 \right] / p^2 m^2 = \text{máximo (II)}$$

Já Saunders, em sua análise, requereu que a curtose (quarto momento sobre o segundo momento quadrado) de todos os carregamentos e suas reflexões fossem um máximo, ou seja

$$q_2 = pm \sum_i \sum_s \ell_{is}^4 / \left(\sum_i \sum_s \ell_{is}^2 \right)^2 = \text{máximo (III)}$$

Ferguson, baseando-se em informações teóricas, denotou simplesmente por

$$q_3 = \sum_i \sum_s \ell_{is}^4 = \text{máximo (IV)}$$

A intenção de todos estes investigadores era a obtenção de uma matriz de fatores com uma máxima tendência de apresentar pequenos e grandes carregamentos. Para fatores ortogonais, os critérios (II), (III) e (IV) são equivalentes pela invariância da soma das comunalidades, $\sum_i \sum_s \ell_{is}^2$, pois este termo, assim como outros constantes, desaparece quando as funções são diferenciadas na busca de pontos críticos. Isso mostra que estes critérios apresentados são equivalentes ao Critério de Carrol para o caso ortogonal. Assim sendo, minimizar (I) é equivalente a maximizar (IV), uma vez que a comunalidade quadrada da variável é

$$\text{constante} = \left(\sum_s \ell_{is}^2 \right)^2 = \sum_s \ell_{is}^4 + 2 \sum_{s < t} \ell_{is}^2 \ell_{it}^2,$$

e a soma das comunalidades quadradas de todas as variáveis é

$$\text{constante} = \sum_i \sum_s \ell_{is}^4 + 2 \sum_i \sum_{s < t} \ell_{is}^2 \ell_{it}^2$$

ou

$$\text{constante} = q_3 + 2f$$

Desta forma, como o Critério Quartimax (q_3) mais duas vezes o Critério de Carrol (f) é uma constante, maximizar q_3 é equivalente a minimizar f .

Na época, Neuhaus e Wrigley perceberam que nenhum destes critérios poderiam ser realisticamente aplicados sem o auxílio de um computador eletrônico, pois os cálculos envolvidos eram muito demorados para computadores mecânicos. Assim, o Método Quartimax foi programado na central de processamento de dados da Universidade de Illinois, que dispunha das máquinas necessárias. Com isso, produziram uma extensa investigação numérica de propriedades empíricas do Método Quartimax.

Os resultados obtidos foram muitas vezes mais audaciosos do que os de Carrol. Sobre a restrição de ortogonalidade do critério de Carrol, ou mesmo do Método Quartimax, não foi mostrado a clara tendência do critério de Carrol, quando a restrição da ortogonalidade é removida. Entretanto, como uma explicação da estrutura simples ortogonal, o Método Quartimax tem uma sistemática parcialidade.

Assim sendo, o Método Quartimax, proposto por Carrol, procura uma melhor interpretação a partir de uma estrutura simples e consiste em se simplificar cada linha, pela maximização na linha da variância dos carregamentos quadrados.

2.4.7.2 Critério Varimax

O Método Varimax de rotação ortogonal foi proposto por Kaiser em 1958. Este nome deve-se ao fato de que o método busca a máxima variância das variáveis. Por isso, o nome Varimax.

Os critérios de rotação anteriormente descritos consideram todos os pm carregamentos simultaneamente. Em cada caso, entretanto, estes critérios podem ser aplicados separadamente para cada linha da matriz de fatores e, pela invariância das comunalidades, somando as linhas para o critério final. Neuhaus e Wrigley definiram como “simplicidade” da composição fatorial da i -ésima variável, a variância dos carregamentos quadrados para esta variável, ou seja

$$q_i^* = \left[m \sum_s (\ell_{is}^2)^2 - \left(\sum_s \ell_{is}^2 \right)^2 \right] / m^2 \quad (\text{V})$$

para obter o critério total para a matriz fatorial inteira, (V) pode então ser somada sobre todas as variáveis resultando assim

$$q^* = \sum_i \left\{ \left[m \sum_s (\ell_{is}^2)^2 - \left(\sum_s \ell_{is}^2 \right)^2 \right] / m^2 \right\} \quad (\text{VI})$$

maximizar q^* é equivalente a maximizar q_3 , novamente pelos termos constantes que desaparecem quando a função é diferenciada.

A equação (VI), muitas vezes produz alguns esclarecimentos no Critério Quartimax. Seu objetivo é simplificar a descrição de cada linha, ou variável da matriz fatorial. E não a simplificação das colunas, ou fatores da matriz fatorial. A implicação disto é que o Critério Quartimax resultará freqüentemente em um fator geral. Sobre a condição (V), não existe motivo pelo qual um grande carregamento para cada variável não possa ocorrer no mesmo fator. Na prática, esta tendência do Critério Quartimax em gerar um fator geral é mais verificada quando a matriz fatorial não rotacionada apresenta um fator geral.

Do ponto de vista da estrutura simples, uma modificação imediata do Critério Quartimax pode ser observada. Definindo, assim, a simplicidade de um fator como a variância de seus carregamentos quadrados,

$$v_s^* = \left[p \sum_i (\ell_{is}^2)^2 - \left(\sum_i \ell_{is}^2 \right)^2 \right] / p^2 \quad (\text{VII})$$

e para o critério em todos os fatores, definindo a máxima simplicidade da matriz fatorial como o máximo de

$$v^* = \sum_s v_s^* = \sum_s \left\{ \left[p \sum_i (\ell_{is}^2)^2 - \left(\sum_i \ell_{is}^2 \right)^2 \right] / p^2 \right\} \quad (\text{VIII})$$

a variância dos carregamentos quadrados das colunas em vez das linhas.

Já que um fator é um vetor de coeficientes de correlação, a maioria das interpretações fatoriais está baseada em coeficientes de correlação que maximizam a interpretação. As correlações que satisfazem esta condição são aquelas cujos coeficientes são ± 1 , os quais indicam um relacionamento funcional, ou zero, que indica um relacionamento não linear. Por outro lado, as medidas médias de correlação são as mais difíceis de serem interpretadas. Assim, pode ser visto porque v_s^* em (VII) pode ser maximizada para o máximo interpretável ou simplicidade de um fator, e mais geralmente, porque a interpretação de toda a matriz fatorial pode ser mais bem considerada quando (VIII) é um máximo.

O critério (VIII) é o original Critério Varimax. Na proposta original deste critério, ele mostra ser matematicamente equivalente, no caso ortogonal, a minimizar

$$c^* = \sum_{s < t} \left\{ \left[p \sum_i \ell_{is}^2 \ell_{it}^2 - \left(\sum_i \ell_{is}^2 \right) \left(\sum_i \ell_{it}^2 \right) \right] / p^2 \right\} \quad (\text{IX})$$

ou seja, minimizando a covariância dos pares de colunas dos carregamentos quadrados e somando sobre todos os possíveis pares de colunas para o critério. O critério (IX) satisfaz a relação análoga para o Critério de Carroll (I), que o Critério Varimax (VIII) faz para o Critério Quartimax (VI).

2.4.7.3 Varimax Normal

Para problemas que utilizam o Critério Quartimax e Varimax, pode-se observar uma tendência a pesos divergentes, que são ligados implicitamente às variáveis por suas communalidades. Assim, uma variável com carregamentos em quatro fatores que tenha comunalidade 0,6, por exemplo, tenderia a influenciar as rotações quatro vezes tanto quanto

outra variável, cuja comunalidade seja 0,3. Assim, quando as raízes quadradas das comunalidades são consideradas no critério, os pesos aparecem e passam a formar um melhor conjunto, pois cada peso tenderia a igualar a influência relativa a cada variável durante a rotação.

Mesmo assim, não parece existir um motivo para a escolha de diferentes esquemas de pesagem. Desta forma, pela complexidade da análise correlacional e com a finalidade de rotação, as variáveis são pesadas igualmente, na intenção de que as dimensões das partes comuns das variáveis tenham mesmo comprimento. O Critério de Varimax poderia então ser rescrito da seguinte forma

$$v = \sum_s \{ [p \sum_i (a_{is}^2 / h_i^2) - [\sum_i (a_{is}^2 / h_i^2)]^2 / p^2] (X) \}$$

onde h_i^2 é a comunalidade da i -ésima variável. Em contraste com (VII) e (VIII), onde a variância das correlações quadradas das variáveis com um fator é maximizada, a variância das correlações quadradas das partes comuns das variáveis (as reflexões das variáveis sobre o espaço de fatores comuns) com um fator está sendo agora maximizada. É importante ressaltar de (X) que não se estão procurando pesos para as variáveis inversamente às raízes quadradas de suas comunalidades. Durante a rotação, estes pesos estendem a parte comum de cada variável para um comprimento unitário, mas depois da rotação, cada uma destas variáveis é reduzida para seu próprio comprimento pela repesagem direta da raiz quadrada da comunalidade da variável.

Os resultados de problemas práticos mostram que o Método Varimax Normal remove efetivamente as pequenas perturbações da solução do Método Varimax original e não parece divergir sistematicamente do que pode ser considerada a melhor estrutura ortogonal simples.

Há uma razão fundamental para tentar se estabelecer o Critério Varimax Normal (X) como uma definição matemática para o problema de rotação. Considere, como exemplo, a

situação ilustrada pela Figura 4. Existem dois grupos de variáveis, cada um dos quais é puro, no sentido de que as reflexões dos vetores das variáveis do grupo sobre o espaço de fatores comuns bidimensional é colinear. Desde que estes grupos estejam localizados produzindo um ângulo menor do que 90 graus, o presente argumento é perfeitamente geral.

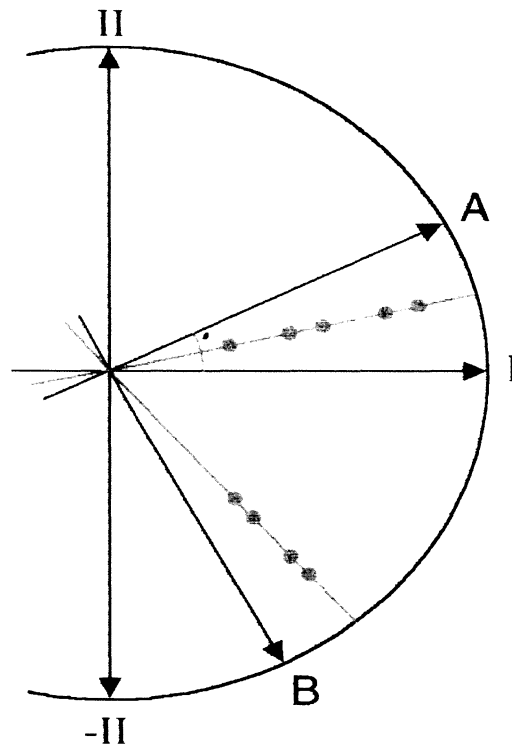


FIGURA 4 - VARIMAX NORMAL

Este é o caso em que uma solução Normal Varimax é invariante sobre alterações na composição das variáveis.

Pode-se mostrar que o ângulo de rotação no plano que maximiza (X) é dado por

$$\phi = \frac{1}{4} \arctan \frac{2 [n \sum_i u_i v_i - \sum_i u_i \sum_i v_i]}{n \sum_i (u_i^2 - v_i^2) - [(\sum_i u_i)^2 - (\sum_i v_i)^2]} \quad (\text{XI})$$

onde

$$u_i = (a_{i1}/h_i)^2 - (a_{i2}/h_i)^2$$

e

$$v_i = 2 (a_{i1}/h_i)(a_{i2}/h_i)$$

Considerando que n_a ($n_a \geq 1$) seja o número de variáveis no primeiro grupo e n_b ($n_b \geq 1$) seja o número de variáveis no segundo grupo ($n = n_a + n_b$), pode-se observar claramente que todas as variáveis do primeiro grupo têm os mesmos valores para u_i e v_i . Considerando que estes valores sejam u_a e v_a , e analogamente que os valores do segundo grupo sejam u_b e v_b , neste caso (XI), reduz-se a

$$\phi = \frac{1}{4} \arctan \frac{2 n_a n_b (u_a v_a + u_b v_b - u_a v_b - u_b v_a)}{n_a n_b (u_a^2 + u_b^2 - v_a^2 - v_b^2 - 2u_a u_b + 2v_a v_b)} \quad (\text{XII})$$

A importância da expressão (XII) é extremamente significativa. O termo de $n_a n_b$ pode ser cancelado indicando que aquele ângulo de rotação não depende do número de variáveis em cada grupo, ou seja, para o caso ilustrado na Figura 4, a solução do Varimax Normal é invariante sobre mudanças na composição das variáveis.

Esta propriedade de invariância parece até ser de maior significado que as tendências numéricas da solução de Varimax Normal para definir matematicamente a doutrina de estrutura simples. Embora a Análise Fatorial pareça ter muitos propósitos, fundamentalmente ela se resume ao seguinte problema: Dado um (infinito) domínio de conteúdo psicológico, deduza a estrutura interna deste domínio com base em uma amostra de n variáveis extraída desse domínio. As possibilidades de sucesso em tais conclusões dependem da extensão com que um fator derivou de uma observação particular ou amostra de variáveis, aproximando os fatores não observáveis no domínio infinito. Se um fator é invariante sobre as variáveis amostrais observadas, ou seja, se existe invariância fatorial, há evidências de que conclusões relativas a fatores do domínio estejam corretas.

A solução do Varimax Normal, de acordo com o resultado anterior, permite tais inferências. Para o problema apresentado na Figura 4, indiferentemente das variáveis amostrais, é possível deduzir precisamente o domínio dos fatores do Método Varimax Normal. Isto não acontece para o Quartimax ou soluções do Varimax original, pois o ângulo de rotação é uma função de n_a e n_b .

É importante ressaltar que o domínio dos fatores do Varimax Normal não é mais significativo que o domínio de algum outro critério. O que se observa é que os fatores encontrados do Varimax Normal terão maior probabilidade de retratar o correspondente domínio real dos fatores.

Embora frequentemente se tenha a impressão de que a estrutura simples é o último critério de um procedimento de rotação, sugere-se que o último critério seja a invariância fatorial. A solução do Varimax Normal foi criada originalmente apenas para atender ao critério de estrutura simples. Mas o fato que mostra matematicamente este tipo de invariância sugere que a idéia de Thurstone fosse direcionada basicamente para a invariância fatorial. O princípio de estrutura simples pode provavelmente ser considerado incidental para a maioria dos conceitos fundamentais de invariância fatorial. Este ponto de vista torna sem sentido os argumentos sobre “realidade psicológica” de fatores gerais, fatores bipolares, fatores de estrutura simples, e outros.

Sem sombra de dúvida, o resultado (XII) é para um caso especial. As correlações entre as variáveis em cada um dos dois grupos têm que formar uma matriz de Spearman perfeita, e a matriz de correlação reduzida deve ser de posto dois. O Varimax Normal, no entanto, apresenta resultados invariantes para dois como grupos de Spearman simultaneamente, e conseqüentemente o Critério do Varimax Normal é uma generalização bidimensional do caso clássico de Spearman. A partir daí, a idéia é generalizar para um caso r -dimensional.

Em geral, na Análise Fatorial, a aplicação do Método Varimax Normal apresenta diferentes carregamentos fatoriais dependendo do número de fatores escolhido. Os resultados podem ser diferentes também dependendo do método de estimação aplicado. Na prática, se existe um fator geral dominante na análise, ou seja, um fator que apresenta elevados carregamentos em praticamente todas as variáveis, sua dominância tenderá a se dispersar e até mesmo desaparecer, após uma rotação fatorial.

2.4.7.4 Rotação Varimax Proposta por Kayser

O Método Varimax de rotação ortogonal procura encontrar eixos com poucos carregamentos altos e com tantos carregamentos baixos ou nulos quanto possível.

Kayser propôs por meio de sua tese de doutoramento intitulada “*The varimax method of factor analysis*” que isto pode ser conseguido, a partir de uma maximização iterativa de uma função de carregamentos quadrados. A descrição do método é feita a seguir.

Seja L uma matriz ($p \times m$) de carregamentos não rotacionados, e seja T uma matriz ortogonal ($m \times m$), a matriz de carregamentos quadrados é $L^* = L.T$, ou seja, ℓ_{ij}^* representa os carregamentos rotacionados da i -ésima variável no j -ésimo fator.

A função ϕ que o critério varimax maximiza é a soma das variâncias dos carregamentos quadrados dentro de cada coluna da matriz de carregamentos, onde cada linha dos carregamentos é normalizada pela comunalidade, ou seja,

$$\phi = \sum_{j=1}^m \sum_{i=1}^p (d_{ij}^2 - \bar{d}_j)^2 = \sum_{j=1}^m \sum_{i=1}^p d_{ij}^4 - p \cdot \sum_{j=1}^m \bar{d}_j^{-2} \quad (I)$$

onde

$$d_{ij} = \frac{\ell_{ij}^*}{h_i} \text{ e } \overline{d_j} = p^{-1} \cdot \sum_{i=1}^p d_{ij}^2$$

Na função ϕ , considerando-se todas as variáveis, $\overline{d_j}$ representa a média dos quadrados dos carregamentos normalizados de cada fator j ; d_{ij}^2 representa todos os quadrados dos carregamentos normalizados. Quando se toma o máximo de ϕ , impõe-se a condição de que o quadrado dos desvios em relação à média de cada fator seja o maior valor possível. Com isso consegue-se carregamentos extremos, ou seja, ou muito altos ou muito baixos. E assim, obtém-se a máxima variância.

O Critério Varimax ϕ é uma função de T , e o algoritmo iterativo encontra uma matriz T que maximiza ϕ .

No caso onde $m=2$ fatores, os cálculos são simplificados. A matriz T dada por

$$T = \begin{bmatrix} \cos \theta & \text{sen } \theta \\ -\text{sen } \theta & \cos \theta \end{bmatrix}$$

representa a rotação dos eixos coordenados no sentido horário de uma ângulo de medida θ .

Assim, considerando-se apenas 2 fatores ($m=2$), a matriz de carregamentos rotacionados ficaria:

$$L_{px2}^* = L_{px2} \cdot T_{2x2}$$

ou seja

$$\begin{bmatrix} \ell_{11}^* & \ell_{12}^* \\ \ell_{21}^* & \ell_{22}^* \\ \vdots & \vdots \\ \ell_{p1}^* & \ell_{p2}^* \end{bmatrix} = \begin{bmatrix} \ell_{11} & \ell_{12} \\ \ell_{21} & \ell_{22} \\ \vdots & \vdots \\ \ell_{p1} & \ell_{p2} \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & \text{sen } \theta \\ -\text{sen } \theta & \cos \theta \end{bmatrix}$$

assim,

$$\ell_{i1}^* = \ell_{i1} \cos \theta - \ell_{i2} \sin \theta \text{ e } \ell_{i2}^* = \ell_{i1} \sin \theta + \ell_{i2} \cos \theta \text{ para } i = 1, 2, \dots, p$$

dividindo pelas raízes quadradas das correspondentes comunalidades, ou seja, normalizando as linhas, tem-se

$$d_{i1} = \frac{\ell_{i1}^*}{h_i} = \frac{\ell_{i1} \cos \theta - \ell_{i2} \sin \theta}{h_i} \text{ e } d_{i2} = \frac{\ell_{i2}^*}{h_i} = \frac{\ell_{i1} \sin \theta + \ell_{i2} \cos \theta}{h_i}, i = 1, 2, \dots, p \text{ (II)}$$

com isso obtém-se a matriz $D(p \times 2)$:

$$D = \left[\left(\frac{\ell_{i1} \cos \theta - \ell_{i2} \sin \theta}{h_i} \right) \quad \left(\frac{\ell_{i1} \sin \theta + \ell_{i2} \cos \theta}{h_i} \right) \right] i = 1, 2, \dots, p.$$

Seja $G_{a,b}$ uma função dos carregamentos não rotacionados e das comunalidades definida por:

$$G_{a,b} = \sum_{i=1}^p \frac{\ell_{i1}^a \cdot \ell_{i2}^b}{h_i^{a+b}} \text{ (III)}$$

Se a equação (I) para $m=2$ é

$$\phi = \sum_{j=1}^2 \sum_{i=1}^p d_{ij}^4 - p \cdot \sum_{j=1}^2 \bar{d}_j^{-2}$$

então,

$$\phi = (d_{11}^4 + d_{21}^4 + \dots + d_{p1}^4) + (d_{12}^4 + d_{22}^4 + \dots + d_{p2}^4) - p \cdot (\bar{d}_1^{-2} + \bar{d}_2^{-2})$$

substituindo (II) em (I), vem

$$\begin{aligned} \phi = & \sum_{i=1}^p \left[\left(\frac{\ell_{i1} \cos \theta - \ell_{i2} \sin \theta}{h_i} \right)^4 + \left(\frac{\ell_{i1} \sin \theta + \ell_{i2} \cos \theta}{h_i} \right)^4 \right] - \\ & - p \cdot \sum_{i=1}^p \left\{ \left[p^{-1} \cdot \left(\frac{\ell_{i1} \cos \theta - \ell_{i2} \sin \theta}{h_i} \right)^2 \right]^2 + \left[p^{-1} \cdot \left(\frac{\ell_{i1} \sin \theta + \ell_{i2} \cos \theta}{h_i} \right)^2 \right]^2 \right\} \text{ (IV)} \end{aligned}$$

Desenvolvendo (IV) , levando-se em conta os resultados 2.17, 2.20 e 2.22, e agrupando em função de (III), a função ϕ ficaria:

$$\phi = \frac{1}{4} [\cos(4\theta).A + \text{sen}(4\theta).B + C]$$

ou

$$4\phi = \cos(4\theta).A + \text{sen}(4\theta).B + C \quad (\text{V})$$

onde

$$A = (G_{0,4} + G_{4,0} - 6G_{2,2} - G_{0,2}^2 - G_{2,0}^2 + 2G_{0,2}G_{2,0} + 4G_{1,1}^2)$$

$$B = 4(G_{1,3} - G_{3,1} - G_{1,1}G_{0,2} + G_{1,1}G_{2,0})$$

$$C = p(3[G_{2,0} + G_{0,2}]^2 - [3G_{0,2}^2 + 3G_{2,0}^2 + 2G_{0,2}G_{2,0} + 4G_{1,1}^2])$$

Multiplicando e dividindo os 2 primeiros termos do segundo membro de (V) por

$(A^2 + B^2)^{1/2} \neq 0$, tem-se:

$$4\phi = (A^2 + B^2)^{1/2} \cdot \left[\cos(4\theta) \cdot \frac{A}{(A^2 + B^2)^{1/2}} + \text{sen}(4\theta) \cdot \frac{B}{(A^2 + B^2)^{1/2}} \right] + C \quad (\text{VI})$$

fazendo $\frac{A}{(A^2 + B^2)^{1/2}} = \cos \alpha$ e $\frac{B}{(A^2 + B^2)^{1/2}} = \text{sen} \alpha$, e substituindo em (VI), vem:

$$4\phi = (A^2 + B^2)^{1/2} [\cos(4\theta) \cdot \cos \alpha + \text{sen}(4\theta) \cdot \text{sen} \alpha] + C$$

usando o resultado 2.16, pode-se escrever

$$4\phi = (A^2 + B^2)^{1/2} \cdot \cos(4\theta - \alpha) + C \quad (\text{VII})$$

O resultado 2.23 mostra que o máximo valor de ϕ é obtido para $4\theta = \alpha$. O valor de α

é obtido de

$$\tan \alpha = \frac{B}{A}$$

e, ainda, para localizar o quadrante, considera-se os sinais do seno e coseno da seguinte forma:

se $A > 0$ e $B > 0$, então α pertence ao 1° quadrante;

se $A < 0$ e $B > 0$, então α pertence ao 2° quadrante;

se $A < 0$ e $B < 0$, então α pertence ao 3° quadrante;

se $A > 0$ e $B < 0$, então α pertence ao 4° quadrante.

No caso em que $m > 2$ (mais do que 2 fatores), uma solução iterativa para a rotação é usada. Nessa solução os fatores são rotacionados aos pares, como no método anteriormente descrito. Assim, o primeiro e segundo fatores são rotacionados de um ângulo θ . O novo primeiro fator é então rotacionado com o terceiro fator original, e assim por diante, até que todos os $\binom{m}{2} = \frac{m(m-1)}{2}$ pares de fatores tenham sido rotacionados. Esta sequência de rotações é chamada ciclo. Desta forma, cada ciclo reúne todos os passos do Método Varimax para 2 fatores apenas. Estes ciclos são repetidos até que se tenha satisfeito algum critério de convergência predeterminado. A convergência se dará aos pares. Quando praticamente não houver mais diferença nos carregamentos de um ciclo para o outro, a solução ótima estará próxima. Pode-se estabelecer também o número de iterações como critério de parada.

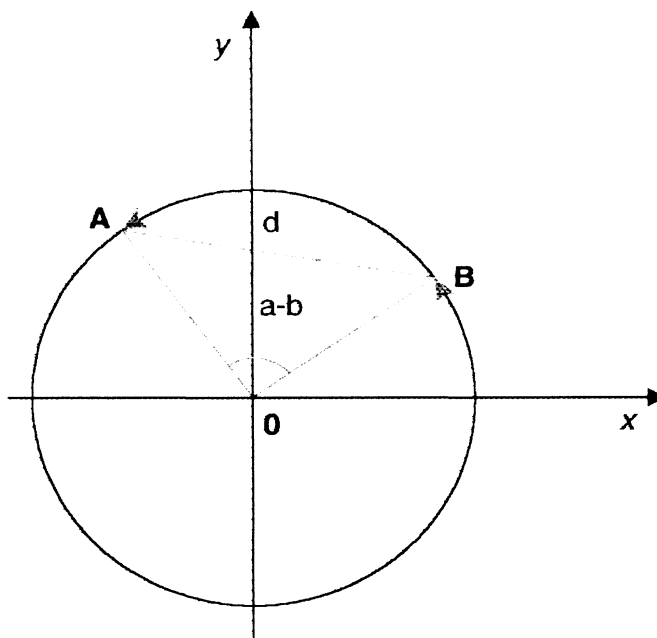
Resultado 2.16:

Para quaisquer arcos a e b , tem-se

$$\cos(a+b) = \cos a \cdot \cos b - \sin a \cdot \sin b$$

Prova:

Sejam A e B as extremidades dos arcos de medidas a e b , respectivamente, então $\widehat{AOB} = a - b$. Sendo d a distância entre A e B , conforme a Figura 5, as coordenadas dos pontos A e B são $A(\cos a, \sin a)$ e $B(\cos b, \sin b)$.

**FIGURA 5 - CICLO TRIGONOMÉTRICO**

Da geometria analítica, utilizando a fórmula da distância entre dois pontos, pode-se escrever:

$$d = d_{AB} = \sqrt{(\cos a - \cos b)^2 + (\sin a - \sin b)^2}$$

$$d^2 = \cos^2 a - 2 \cos a \cos b + \cos^2 b + \sin^2 a - 2 \sin a \sin b + \sin^2 b$$

$$d^2 = \left(\underbrace{\sin^2 a + \cos^2 a}_1 \right) + \left(\underbrace{\sin^2 b + \cos^2 b}_1 \right) - 2(\cos a \cos b + \sin a \sin b)$$

$$d^2 = 2 - 2(\cos a \cos b + \sin a \sin b) \quad (\text{I})$$

Pela lei dos cossenos no triângulo AOB , vem:

$$AB^2 = AO^2 + BO^2 - 2 \cdot AO \cdot BO \cdot \cos(\widehat{AOB})$$

considerando a circunferência unitária:

$$d^2 = 1^2 + 1^2 - 2 \cdot 1 \cdot 1 \cdot \cos(a - b) \quad (\text{II})$$

De (I) e (II), tem-se

$$2 - 2 \cdot \cos(a - b) = 2 - 2 \cdot (\cos a \cdot \cos b + \operatorname{sen} a \cdot \operatorname{sen} b)$$

logo

$$\cos(a - b) = \cos a \cdot \cos b + \operatorname{sen} a \cdot \operatorname{sen} b$$

Resultado 2.17:

Para quaisquer arcos a e b , tem-se

$$\cos(a + b) = \cos a \cdot \cos b - \operatorname{sen} a \cdot \operatorname{sen} b$$

Prova :

Lembrando que $\cos(-x) = \cos x$ e $\operatorname{sen}(-x) = -\operatorname{sen} x$, tem-se

$$\cos(a + b) = \cos[a - (-b)] = \cos a \cdot \cos(-b) + \operatorname{sen} a \cdot \operatorname{sen}(-b)$$

logo

$$\cos(a + b) = \cos a \cdot \cos b - \operatorname{sen} a \cdot \operatorname{sen} b$$

Resultado 2.18:

Qualquer que seja o ângulo θ , vale a relação

$$\cos(2\theta) = \cos^2 \theta - \operatorname{sen}^2 \theta$$

Prova:

Do resultado 2.17, fazendo $a = b = \theta$, vem

$$\cos(\theta + \theta) = \cos\theta \cdot \cos\theta - \operatorname{sen}\theta \cdot \operatorname{sen}\theta$$

assim,

$$\cos(2\theta) = \cos^2 \theta - \operatorname{sen}^2 \theta$$

Resultado 2.19:

Para quaisquer arcos a e b , tem-se

$$\operatorname{sen}(a+b) = \operatorname{sen}a \cdot \cos b + \operatorname{sen}b \cdot \operatorname{sen}a$$

Prova:

Partindo da igualdade entre seno e coseno para arcos complementares:

$$\operatorname{sen}(a+b) = \cos\left[\frac{\pi}{2} - (a+b)\right] = \cos\left[\left(\frac{\pi}{2} - a\right) - b\right]$$

com o resultado 2.17, pode-se desenvolver o 2º membro

$$\operatorname{sen}(a+b) = \cos\left(\frac{\pi}{2} - a\right) \cdot \cos b + \operatorname{sen}\left(\frac{\pi}{2} - a\right) \operatorname{sen} b$$

logo

$$\operatorname{sen}(a + b) = \operatorname{sen}a \cdot \operatorname{cos}b + \operatorname{sen}b \cdot \operatorname{cos}a$$

Resultado 2.20:

Qualquer que seja o ângulo θ , vale a relação

$$\operatorname{sen}(2\theta) = 2\operatorname{sen}\theta \cdot \operatorname{cos}\theta$$

Prova:

Do resultado 2.21, fazendo $a = b = \theta$, vem

$$\operatorname{sen}(\theta + \theta) = \operatorname{sen}\theta \cdot \operatorname{cos}\theta + \operatorname{sen}\theta \cdot \operatorname{cos}\theta$$

assim

$$\operatorname{sen}(2\theta) = 2\operatorname{sen}\theta \cdot \operatorname{cos}\theta$$

Resultado 2.23 :

Para todo ângulo θ , vale a relação

$$\operatorname{sen}\left(\frac{\theta}{2}\right) = \pm \sqrt{\frac{1 - \operatorname{cos}\theta}{2}}$$

Prova:

Do resultado 2.17, pode-se escrever

$$\operatorname{cos}(2x) = \operatorname{cos}^2 x - \operatorname{sen}^2 x \quad (\text{I})$$

Da relação fundamental da trigonometria sabe-se que:

$$\cos^2 x = 1 - \operatorname{sen}^2 x \quad (\text{II})$$

substituindo (II) em (I), vem

$$\cos(2x) = 1 - 2\operatorname{sen}^2 x$$

isolando na última equação $\operatorname{sen} x$, tem-se

$$\operatorname{sen} x = \pm \sqrt{\frac{1 - \cos(2x)}{2}}$$

fazendo $x = \frac{\theta}{2}$ chega-se a

$$\operatorname{sen}\left(\frac{\theta}{2}\right) = \pm \sqrt{\frac{1 - \cos\theta}{2}}$$

Resultado 2.22:

Qualquer que seja o ângulo θ , vale a relação

$$4(\cos^4 \theta + \operatorname{sen}^4 \theta) = 3 + \cos(4\theta)$$

Prova:

Do resultado 2.17, pode-se escrever:

$$\cos(4\theta) = \cos^2(2\theta) - \operatorname{sen}^2(2\theta)$$

novamente, dos resultados 2.17 e 2.20, tem-se

$$\cos(4\theta) = (\cos^2 \theta - \operatorname{sen}^2 \theta)^2 - (2 \operatorname{sen} \theta \cos \theta)^2$$

desenvolvendo

$$\cos(4\theta) = \cos^4 \theta - 2\sin^2 \theta \cos^2 \theta + \sin^4 \theta - 4\sin^2 \theta \cos^2 \theta$$

$$\cos(4\theta) = \cos^4 \theta + \sin^4 \theta - 6\sin^2 \theta \cos^2 \theta$$

levando em conta que $6\sin^2 \theta \cos^2 \theta = \frac{3}{2} \cdot \sin^2(2\theta)$, tem-se

$$\cos(4\theta) = \cos^4 \theta + \sin^4 \theta - \frac{3}{2} \cdot \sin^2(2\theta) \quad (\text{I})$$

do resultado 2.21, pode-se escrever

$$\sin\left(\frac{\alpha}{2}\right) = \pm \sqrt{\frac{1 - \cos \alpha}{2}}$$

fazendo $\alpha = 4\theta$, vem

$$\sin(2\theta) = \pm \sqrt{\frac{1 - \cos(4\theta)}{2}}$$

quadrando ambos os membros :

$$\sin^2(2\theta) = \frac{1 - \cos(4\theta)}{2} \quad (\text{II})$$

Substituindo (II) em (I),vem:

$$\cos(4\theta) = \cos^4 \theta + \sin^4 \theta - \frac{3}{2} \left[\frac{1 - \cos(4\theta)}{2} \right]$$

$$4 \cdot \cos(4\theta) = 4\cos^4 \theta + 4\sin^4 \theta - 3 + 3\cos(4\theta)$$

Então

$$4(\cos^4 \theta + \sin^4 \theta) = 3 + \cos(4\theta)$$

Resultado 2.23:

A função ϕ tem um máximo para $\alpha = 4\theta$.

Prova :

A função ϕ é dada por:

$$4\phi = (A^2 + B^2)^{1/2} \cdot \cos(4\theta - \alpha) + C$$

ou

$$\phi = \frac{1}{4} \cdot (A^2 + B^2)^{1/2} \cdot \cos(4\theta - \alpha) + \frac{1}{4} \cdot C$$

derivando a função ϕ , em relação a θ :

$$\frac{\partial \phi}{\partial \theta} = -\frac{1}{4} \cdot (A^2 + B^2)^{1/2} \cdot \text{sen}(4\theta - \alpha) \cdot 4 + 0$$

$$\frac{\partial \phi}{\partial \theta} = -(A^2 + B^2)^{1/2} \cdot \text{sen}(4\theta - \alpha)$$

Igualando a zero a derivada primeira encontram-se os pontos críticos:

$$\frac{\partial \phi}{\partial \theta} = 0 \Rightarrow -(A^2 + B^2)^{1/2} \cdot \text{sen}(4\theta - \alpha) = 0$$

Se $A \neq 0$ e $B \neq 0$, então $\text{sen}(4\theta - \alpha) = 0$

$$4\theta - \alpha = k\pi, \quad k \in \mathbb{Z}$$

então

$$\alpha = 4\theta - k\pi$$

Em particular, para $k = 0$, obtém-se um ponto crítico, logo:

$$k = 0 \Rightarrow \alpha = 4\theta$$

Para a verificação do ponto crítico, calcula-se a derivada segunda da função ϕ :

$$\frac{\partial^2 \phi}{\partial \theta^2} = -4 \cdot (A^2 + B^2)^{1/2} \cdot \cos(4\theta - \alpha)$$

Substituindo $\alpha = 4\theta$ na equação anterior, vem:

$$-4.(A^2 + B^2)^{1/2} \cdot \cos(4\theta - 4\theta) = -4.(A^2 + B^2)^{1/2} < 0$$

Portanto, se a derivada segunda da função é negativa, então $\alpha = 4\theta$ é um ponto de máximo da função ϕ .

2.4.7.5 Rotação Varimax Proposta por Horst

No método original de Kayser, proposto em 1958, os fatores são rotacionados em pares até que os carregamentos converjam para os valores finais. Logo após, um outro método foi discutido por Horst em “*Factor Analysis of Data Matrices*”, no ano de 1965. O procedimento de Horst, que ele intitulou “os fatores simultâneos da solução varimax” é descrito a seguir.

A idéia deste procedimento é de que, em cada iteração, os fatores converjam simultaneamente. Isto traz vantagens em relação ao método anterior não só pelo tempo computacional, mas também pelo menor acúmulo de erros de arredondamento.

Para uma melhor compreensão do procedimento, inicialmente são consideradas explicitamente as seguintes matrizes:

a) Matriz de carregamentos não rotacionados

$$L_{pxm} = \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1m} \\ l_{21} & l_{22} & \dots & l_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ l_{i1} & l_{i2} & \dots & l_{im} \\ \cdot & \cdot & \cdot & \cdot \\ l_{p1} & l_{p2} & \dots & l_{pm} \end{bmatrix}$$

b) Matriz ortogonal de rotação

$$T_{mxm} = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1r} & \dots & t_{1k} \\ t_{21} & t_{22} & \dots & t_{2r} & \dots & t_{2k} \\ \cdot & \cdot & \dots & \cdot & \dots & \cdot \\ t_{k1} & t_{k2} & \dots & t_{kr} & \dots & t_{kk} \end{bmatrix}$$

c) Matriz de carregamentos rotacionados

$$L_{pxm}^* = \begin{bmatrix} l_{11}^* & l_{12}^* & \dots & l_{1k}^* \\ l_{21}^* & l_{22}^* & \dots & l_{2k}^* \\ \cdot & \cdot & \cdot & \cdot \\ l_{i1}^* & l_{i2}^* & \dots & l_{ik}^* \\ \cdot & \cdot & \cdot & \cdot \\ l_{p1}^* & l_{p2}^* & \dots & l_{pk}^* \end{bmatrix}$$

Sendo $l_i' = (l_{i1}, l_{i2}, \dots, l_{im})$ e $t_r' = (t_{1r}, t_{2r}, \dots, t_{mr})$, respectivamente a i -ésima linha da matriz L e a r -ésima coluna da matriz T , tem-se:

$$l_{ir}^* = l_i' t_r = l_{i1} t_{1r} + l_{i2} t_{2r} + \dots + l_{im} t_{mr}$$

Sejam os escalares

$$d_r = \sum_{i=1}^p (l_{ir}^*)^2 = \sum_{i=1}^p (l_i' t_r)^2, \quad r = 1, 2, \dots, m$$

ou seja, d_r é a soma dos quadrados dos carregamentos da r -ésima coluna de L^* .

O critério x que é maximizado no Método Varimax simultâneo é dado por

$$x = \sum_{r=1}^m \left[\sum_{i=1}^p (\ell_{ir}^2 * -d_r / p)^2 \right] = \sum_r \left[\sum_i \ell_{ir}^4 * -d_r^2 / p \right] = \sum_r \left[\sum_i (\ell_i' t_r)^4 - d_r^2 / p \right]$$

onde x representa a soma dos quadrados dos desvios dos valores de $\ell_{ir}^2 *$, cada medida como um desvio da correspondente média da coluna d_r / p .

A maximização é em relação aos elementos da matriz T . As colunas de T satisfazem as condições:

$$\begin{aligned} t_r' t_r &= 1 \\ t_r' t_s &= 0 \quad (r \neq s) \end{aligned}$$

Deve-se igualar a zero as derivadas em relação aos elementos de T da expressão

$$y = x - \sum_r \sum_s a_{rs} t_r' t_s$$

onde os coeficientes a são multiplicadores indeterminados, com $a_{rs} = a_{sr}$.

Substituindo a expressão de x na expressão anterior tem-se:

$$y = \sum_r \left[\sum_i (\ell_i' t_r)^4 - \left(\sum_i (\ell_i' t_r)^2 \right)^2 / p \right] - 2 \sum_r \sum_s a_{rs} t_r' t_s$$

diferenciando em relação a t , vem:

$$\frac{\partial y}{\partial t_s} = 4 \cdot \sum_i [(\ell_i' t_s)^3 \cdot \ell_i] - (4 \cdot d_s / p) \cdot \sum_i [(\ell_i' t_s) \cdot \ell_i] - 4 \cdot \sum_r (a_{rs} \cdot t_r)$$

ou

$$\frac{\partial y}{\partial t_s} = 4 \cdot \sum_i (c_{is} - d_s \ell_{is}^* / p) \ell_i - 4 \cdot \sum_r (a_{rs} \cdot t_r), \text{ com } c_{ir} = (\ell_i' t_r)^3 = \ell_{ir}^3 *$$

A expressão $\frac{\partial y}{\partial t_s}$ é a s -ésima coluna da matriz

$$4.[L'C - (1/p)L'L^*.D - TA]$$

onde

C_{pxm} é uma matriz cujos elementos são c_{ir} ;

A_{mxm} é simétrica com elementos a_{rs} ;

D_{mxm} é uma matriz diagonal com elementos d_1, d_2, \dots, d_m .

Portanto, derivando y em relação a todas as colunas de T , ou seja, $\frac{\partial y}{\partial t_i}$, $i = 1, 2, \dots, m$,

tem-se:

$$\frac{\partial y}{\partial T} = 4 \left(\underbrace{L'[C - (1/p)L^*.D]}_B - TA \right)$$

ou

$$\frac{\partial y}{\partial T} = 4.(B - TA)$$

Igualando a zero a derivada de y em relação a T , encontra-se a matriz ortogonal T que maximiza x satisfazendo a equação

$$TA = B$$

Pode-se mostrar que se A é positiva definida e simétrica, a equação anterior corresponde ao máximo para x . Pré-multiplicando por T^o , tem-se:

$$T^o TA = T^o B$$

Como T é ortogonal, $T^o T = I$, então

$$A = T^o B$$

Substituindo a expressão da matriz B na equação anterior, vem

$$A = T^o L^{*o} . (C - (1/p)L^*.D)$$

Lembrando que $L^* = L.T$, tem-se

$$A = L^{*'} \cdot (C - (1/p) \cdot L^* \cdot D)$$

ou

$$A = L^{*'} \cdot C - (1/p) \cdot L^{*'} \cdot L^* \cdot D$$

Assim, $L^{*'} \cdot C - (1/p) \cdot L^{*'} \cdot L^* \cdot D$ é simétrica quando x é maximizado.

O r -ésimo elemento da diagonal de A é

$$\sum_i \ell_{ir}^4 \cdot (1/p) \cdot \left(\sum_i \ell_{ir}^2 \right)^2$$

Da definição de x , isto mostra que $\text{tr}(A)$ é seu valor máximo.

As matrizes T , A e B que satisfazem

$$TA = B$$

são encontradas por um procedimento iterativo. Começa-se com uma aproximação inicial T_1 .

Isto leva a uma aproximação $L_1 = L_0 T_1$ para L^* , e de L_1 obtém-se aproximações C_1 , D_1 e B_1 de C , D e B , respectivamente. Procura-se, então, encontrar uma matriz simétrica e positiva definida A_1 e uma matriz ortogonal T_2 que satisfaçam a equação

$$T_2 A_1 = B_1$$

Observa-se que pré-multiplicando cada lado pela transposta de B_1 , tem-se

$$B_1' T_2 A_1 = B_1' B_1 \Rightarrow A_1 A_1 = B_1' B_1 \Rightarrow A_1^2 = B_1' B_1$$

Agora $B_1' B_1$ é simétrica e positiva definida de ordem m .

Portanto, pode-se expressá-la, através da decomposição espectral, na forma

$$U \Delta^2 U'$$

onde U é a matriz ortogonal, cujas colunas são os autovetores coluna normalizados de A_1^2 e Δ

é a matriz diagonal, cujos elementos diagonais positivos são os autovalores também de A_1^2 .

Desta forma, A_1 será dada por:

$$A_1 = U\Delta U'$$

Como $T_2 A_1 = B_1$, então $T_2 = B_1 A_1^{-1}$.

O procedimento é repetido com T_2 no lugar de T_1 e, se necessário, com uma seqüência de matrizes T_s . Pode-se mostrar que para T_1 suficientemente próxima de T , a seqüência converge para T . A correspondente seqüência de matrizes A_s que tem a forma dos valores de $\text{tr}(A)$ cresce convergindo para o critério máximo de x . Na prática, uma boa aproximação inicial de T não é necessária, logo toma-se usualmente $T_1 = I$ e $L_1 = L_0$.

É importante ressaltar que, segundo Horst (1965), antes de se iniciar o procedimento de maximização, a matriz de carregamentos L deve apresentar suas linhas padronizadas. Desta forma, cada vetor linha da matriz L deve ter módulo unitário. Concluída a maximização, após a matriz T de rotação ter sido obtida, as linhas são redimensionadas para que as communalidades de cada variável assumam o valor correto. Como visto anteriormente, as communalidades ficam inalteradas mediante uma rotação fatorial.

O Método Varimax apresenta duas limitações principais. Em primeiro lugar, os carregamentos obtidos da rotação fatorial podem se apresentar completamente diferentes, caso fatores adicionais sejam considerados na rotação. É muito importante, portanto, que a opção do número de fatores seja coerente com a análise que se deseja realizar. Em segundo lugar, o método não se mostra muito eficiente nos casos em que existe um fator geral dominante. Um fator dominante tende a desaparecer após uma rotação de fatores. Ocorrendo, assim, certa perda de informação.

2.4.7.6 Método Promax

O Método Promax foi desenvolvido por HENDRICKSON e WHITE (LAWLEY e MAXWELL, 1971, p. 76) em "*Promax: a quick method for rotation to oblique simple structure*", em 1964 a partir de idéias de Thurstone, Carrol e Kayser.

Em muitos casos, os carregamentos fatoriais podem apresentar uma interpretação simplificada através de uma transformação oblíqua, na qual os fatores são correlacionados. Para se obter realmente uma interpretação melhor na forma de uma estrutura simples, os coeficientes de correlação devem ser calculados e levados em consideração em qualquer interpretação dos resultados.

Este método se inicia com a rotação varimax dos carregamentos originais. Em vez de apenas considerar a matriz de carregamentos L de ordem pxm , constrói-se outra matriz Q , de mesma ordem, cujos elementos são definidos por

$$q_{ij} = |\ell_{ij}^{r-1}| \ell_{ij}$$

onde r é um inteiro tal que $r > 1$. Assim, $q_{ij} = 0$ se $\ell_{ij} = 0$; caso contrário, q_{ij} tem o mesmo sinal de ℓ_{ij} e tem o mesmo valor absoluto de ℓ_{ij}^r . Procura-se uma matriz U , não necessariamente ortogonal, tal que para $j = 1, 2, \dots, m$, a j -ésima coluna de LU é um máximo em relação à j -ésima coluna de Q . Quando U é encontrada, tem-se como consequência, em geral, um aumento na magnitude absoluta dos carregamentos maiores e diminuição nos carregamentos que são pequenos. O valor escolhido para r é muito subjetivo. É mais uma questão de se avaliar com tentativas. Entretanto, experiências mostram que valores de r maiores que 4, mesmo que simplifiquem a interpretação dos fatores, fazem com que estes fatores sejam altamente correlacionados, o que na prática não é muito interessante.

Sejam q_j e u_j a j -ésima coluna de Q e U , respectivamente. Então, para cada valor de j , escolhe-se u_j , de forma a minimizar a expressão

$$(q_j - Lu_j)'(q_j - Lu_j)$$

A derivada da expressão em relação ao vetor u_j é

$$-2L'(q_j - Lu_j)$$

Igualando a zero, tem-se

$$q_j - Lu_j = 0$$

ou

$$Lu_j = q_j$$

Pré-multiplicando por L' , vem

$$(L'L)u_j = L'q_j$$

Levando-se em consideração todos os valores de j , tem-se

$$(L'L)U = L'Q$$

ou

$$U = (L'L)^{-1}L'Q$$

Na prática é conveniente padronizar as colunas de U para que os fatores transformados estejam contabilizados nas mesmas unidades de variância. Isto pode ser obtido encontrando uma matriz diagonal D , com elementos diagonais positivos tal que

$$D^2 = \text{diag}[(U'U)^{-1}]$$

A transformação ocorre com a troca da matriz U , pela matriz M , através da relação

$$M = UD$$

e a matriz de carregamentos transformados L^* é dada por

$$L^* = LM$$

Assim, tem-se

$$LL' = L'M^{-1}M^{-1}L'$$

$$LL' = L'(MM)^{-1}L'$$

$$LL' = L'\Phi L'$$

onde

$$\Phi = (MM)^{-1} = D^{-1}(UU)^{-1}D^{-1}$$

é a matriz de covariância dos fatores transformados. Além disso, Φ tem elementos diagonais unitários. Conseqüentemente, é também a matriz de correlação para os novos fatores, os quais foram de fato padronizados.

Em geral, na rotação de fatores, utiliza-se a rotação Varimax normal, considerando-se os fatores não correlacionados. Mesmo assim, na prática, ótimos resultados têm sido obtidos com os fatores correlacionados através do Método Promax e similares. Uma rotação oblíqua é freqüentemente utilizada como ajuda adicional na Análise Fatorial.

Graficamente, uma rotação oblíqua para uma estrutura simples corresponde a uma rotação não rígida das coordenadas do sistema, de modo que os eixos rotacionados passem muito próximos dos grupos das variáveis. Uma rotação oblíqua procura expressar cada variável em termos de um número mínimo de fatores, preferivelmente um fator simples.

2.4.8 ESCORES FATORIAIS

Na Análise Fatorial, o interesse está usualmente centrado nos parâmetros do modelo. Entretanto, em determinados problemas, os valores estimados dos fatores comuns, denominados escores fatoriais, podem também ser requeridos. Essas quantidades são

freqüentemente usadas para propósitos de diagnósticos como também para introduzir uma análise posterior.

Escores fatoriais são estimativas dos valores para os vetores fatoriais aleatórios não observáveis $F_j, j = 1, 2, \dots, n$. Assim, os escores fatoriais

\hat{f}_j são estimativas dos valores f_j , obtidos de F_j (j -ésimo caso)

A estimação é complicada pelo fato de que as quantidades f_j e ε_j não são observadas como x_j . Para resolver este problema, muitas heurísticas foram debatidas, e várias aproximações têm sido obtidas da estimação dos valores fatoriais. Nesta dissertação, três métodos de estimação serão descritos: o Método dos Mínimos Quadrados Ponderados, o Método da Regressão e o Método da Regressão para Fatores Correlacionados.

Todos os três métodos têm dois pontos em comum na estimação:

- a) tratam os carregamentos fatoriais estimados $\hat{\ell}_{ij}$, e as variâncias específicas $\hat{\psi}_i$, como se fossem os verdadeiros valores;
- b) envolvem transformações lineares dos dados originais, muitas vezes centrados ou padronizados.

Tipicamente, os carregamentos fatoriais rotacionados estimados são mais usados do que os carregamentos fatoriais não rotacionados. As fórmulas apresentadas aqui não diferenciam o uso de cada um. Assim sendo, tanto os carregamentos fatoriais rotacionados quanto os não rotacionados podem ser usados. Apresenta-se a seguir o Método dos Mínimos Quadrados Ponderados.

2.4.8.1 Método dos Mínimos Quadrados Ponderados

Para este modelo, supõe-se primeiramente que a média μ , a matriz de carregamentos fatoriais L , e a matriz de variância específica Ψ , são conhecidas do modelo fatorial

$$\underset{(px1)}{\tilde{X}} - \underset{(px1)}{\tilde{\mu}} = \underset{(pxm)}{\tilde{L}} \cdot \underset{(mx1)}{\tilde{F}} + \underset{(px1)}{\tilde{\varepsilon}}$$

Considera-se ainda que os fatores específicos $\varepsilon' = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p]$ são erros. Desde que $Var(\varepsilon_i)$ não precise ser igual a ψ_i , $i = 1, 2, \dots, p$, BARTLETT (citado por JOHNSON e WICHERN, 1992) sugeriu que os valores fatoriais comuns sejam estimados pelo Método dos Mínimos Quadrados Ponderados.

A soma dos erros quadrados, ponderada pelos inversos de suas variâncias, é dada por

$$\sum_{i=1}^p \frac{\varepsilon_i^2}{\psi_i} = \underset{(1 \times p)}{\tilde{\varepsilon}}' \underset{(p \times p)}{\Psi}^{-1} \underset{(p \times 1)}{\tilde{\varepsilon}} = (\underset{(1 \times p)}{\tilde{x}} - \underset{(1 \times m)}{\tilde{\mu}} - \underset{(p \times m)}{\tilde{L}} \underset{(m \times 1)}{\tilde{f}})' \underset{(p \times p)}{\Psi}^{-1} (\underset{(1 \times p)}{\tilde{x}} - \underset{(1 \times p)}{\tilde{\mu}} - \underset{(p \times m)}{\tilde{L}} \underset{(m \times 1)}{\tilde{f}})$$

A proposta é a escolha de estimadores \hat{f} no intuito de minimizar a expressão anterior.

O resultado 2.24 mostra que a solução para o estimador do Método de Mínimos Quadrados Ponderados é

$$\hat{f} = (\hat{L}' \hat{\Psi}^{-1} \hat{L})^{-1} \hat{L}' \hat{\Psi}^{-1} (\underset{(1 \times p)}{\tilde{x}} - \underset{(1 \times p)}{\tilde{\mu}})$$

Tomando \hat{L} , $\hat{\Psi}$ e $\hat{\mu} = \bar{x}$ como valores verdadeiros, o j -ésimo caso fica:

$$\hat{f}_j = \left(\hat{L}' \hat{\Psi}^{-1} \hat{L} \right)^{-1} \hat{L}' \hat{\Psi}^{-1} (\underset{(1 \times p)}{\tilde{x}}_j - \bar{x})$$

Quando \hat{L} e $\hat{\Psi}$ são determinados pelo Método da Máxima Verossimilhança, estas estimativas devem satisfazer a condição de singularidade, isto é, $\hat{L}'\hat{\Psi}^{-1}\hat{L} = \hat{\Delta}$ deve ser uma matriz diagonal.

Resultado 2.24:

O estimador do Método de Mínimos Quadrados Ponderados é dado por

$$\hat{f} = (L'\Psi^{-1}L)^{-1}L'\Psi^{-1}(x-\mu)$$

Prova:

A soma dos erros quadrados ponderados é dada por

$$\sum_{i=1}^p \frac{\varepsilon_i^2}{\psi_i} = \varepsilon'\Psi^{-1}\varepsilon = (x-\mu-Lf)'\Psi^{-1}(x-\mu-Lf)$$

Como o menor erro quadrado é zero, deve-se igualar a zero a última parte. Assim,

$$(x-\mu-Lf)'\Psi^{-1}(x-\mu-Lf) = 0$$

A nulidade só ocorre se:

$$x-\mu-Lf = 0$$

logo

$$Lf = x-\mu$$

$$L^{-1} \left(\hat{L} \hat{f} \right) = L^{-1} \cdot (\underline{x} - \underline{\mu})$$

$$\hat{f} = L^{-1} \cdot (\underline{x} - \underline{\mu})$$

$$\hat{f} = L^{-1} \cdot (\Psi \cdot \Psi^{-1}) (\underline{x} - \underline{\mu})$$

$$\hat{f} = L^{-1} \cdot (\Psi [(L')^{-1} L'] \Psi^{-1}) (\underline{x} - \underline{\mu})$$

pela associatividade do produto matricial, pode-se escrever :

$$\hat{f} = [L^{-1} \cdot \Psi (L')^{-1}] L' \Psi^{-1} \cdot (\underline{x} - \underline{\mu})$$

Portanto,

$$\hat{f} = (L' \Psi^{-1} L)^{-1} L' \Psi^{-1} \cdot (\underline{x} - \underline{\mu})$$

é a solução para o estimador do Método de Mínimos Quadrados Ponderados.

Sendo assim, para carregamentos estimados pelo Método da Máxima Verossimilhança, pode-se montar o seguinte Modelo de Mínimos Quadrados Ponderados:

$$\hat{f}_{\sim j} = \left(\hat{L}' \hat{\Psi}^{-1} \hat{L} \right)^{-1} \hat{L}' \hat{\Psi}^{-1} (\underline{x}_{\sim j} - \underline{\mu}) = \hat{\Delta}^{-1} \hat{L}' \hat{\Psi}^{-1} (\underline{x}_{\sim j} - \bar{\underline{x}}), \quad j = 1, 2, \dots, n$$

ou, se a matriz de correlação é fatorada,

$$\hat{f}_{\sim j} = \left(\hat{L}'_z \hat{\Psi}_z^{-1} \hat{L}_z \right)^{-1} \hat{L}'_z \hat{\Psi}_z^{-1} \underline{z}_{\sim j} = \hat{\Delta}_z^{-1} \hat{L}'_z \hat{\Psi}_z^{-1} \underline{z}_{\sim j}, \quad j = 1, 2, \dots, n$$

no qual $\underline{z}_{\sim j} = D^{-1/2} (\underline{x}_{\sim j} - \bar{\underline{x}})$ e $\hat{\rho} = \hat{L}'_z \hat{L}_z + \hat{\Psi}_z$.

Os escores anteriores poderiam também ser obtidos após uma rotação de fatores. Em

vez de $\hat{f}_{\sim j}$ seria $\hat{f}_{\sim j}^*$, $j = 1, 2, \dots, n$.

Caso os carregamentos fatoriais fossem estimados pelo Método de Componentes Principais, para generalizar os escores fatoriais, é comum o uso de um procedimento de mínimos quadrados sem ponderação. Implicitamente assume-se que ψ_i é igual ou quase igual ao valor verdadeiro. Os escores fatoriais seriam, então, dados por

$$\hat{f}_{\sim j} = \left(\hat{L}' \hat{L} \right)^{-1} \hat{L}' (\hat{x}_{\sim j} - \bar{x}) \quad \text{ou, para dados padronizados} \quad \hat{f}_{\sim j} = \left(\hat{L}_z' \hat{L}_z \right)^{-1} \hat{L}_z' \hat{z}_{\sim j}.$$

Desde que

$$L = \left[\sqrt{\hat{\lambda}_1} \cdot \hat{e}_{\sim 1} \quad \sqrt{\hat{\lambda}_2} \cdot \hat{e}_{\sim 2} \quad \dots \quad \sqrt{\hat{\lambda}_m} \cdot \hat{e}_{\sim m} \right]$$

tem-se

$$\hat{f}_{\sim j} = \begin{bmatrix} \frac{1}{\sqrt{\hat{\lambda}_1}} \hat{e}_{\sim 1}' (\hat{x}_{\sim j} - \bar{x}) \\ \frac{1}{\sqrt{\hat{\lambda}_2}} \hat{e}_{\sim 2}' (\hat{x}_{\sim j} - \bar{x}) \\ \vdots \\ \frac{1}{\sqrt{\hat{\lambda}_m}} \hat{e}_{\sim m}' (\hat{x}_{\sim j} - \bar{x}) \end{bmatrix}$$

Para estes escores fatoriais,

$$\frac{1}{n} \sum_{j=1}^n \hat{f}_{\sim j} = 0 \quad (\text{média amostral})$$

e

$$\frac{1}{n-1} \sum_{j=1}^n \hat{f}_{\sim j} \hat{f}_{\sim j}' = I \quad (\text{covariância amostral})$$

2.4.8.2 Método da Regressão

Considerando o Modelo Fatorial Ortogonal $\tilde{X} - \tilde{\mu} = L\tilde{F} + \tilde{\varepsilon}$, neste método, novamente as matrizes de carregamentos L e de variâncias específicas Ψ serão tratadas como conhecidas. Quando os fatores comuns \tilde{F} e os fatores específicos (ou erros) $\tilde{\varepsilon}$ são conjuntamente normalmente distribuídos com as suposições do Modelo Fatorial Ortogonal, a combinação linear $\tilde{X} - \tilde{\mu} = L\tilde{F} + \tilde{\varepsilon}$ tem uma distribuição $N_p(0, LL' + \Psi)$. O resultado 2.28 comprova isso. Além disso, a distribuição conjunta de $(\tilde{X} - \tilde{\mu})$ e \tilde{F} é $N_{m+p}(0, \Sigma^*)$, na qual

$$\Sigma^*_{(m+p) \times (m+p)} = \begin{bmatrix} \Sigma = LL' + \Psi & \vdots & L \\ \text{.....} & \vdots & \text{.....} \\ L' & \vdots & I \end{bmatrix}$$

(pxp) (pxm)
 (mxp) (mxm)

e 0 é um vetor nulo de dimensão (mxp) . Usando o resultado 2.29, a distribuição condicional de $\tilde{F} | \tilde{x}$ é multivariada normal com

$$\text{Média} = E(\tilde{F} | \tilde{x}) = L' \Sigma^{-1} (\tilde{x} - \tilde{\mu}) = L' (LL' + \Psi)^{-1} (\tilde{x} - \tilde{\mu}) \quad (\text{I})$$

e

$$\text{Covariância} = \text{Cov}(\tilde{F} | \tilde{x}) = I - L' \Sigma^{-1} L = I - L' (LL' + \Psi)^{-1} L \quad (\text{II})$$

As quantidades $L(LL' + \Psi)^{-1}$ de (I) são os coeficientes na regressão multivariada dos fatores nas variáveis. Estimativas destes coeficientes produzem escores fatoriais que são análogos às estimativas dos valores médios condicionados na análise de regressão multivariada. Conseqüentemente, dado qualquer vetor de observações \tilde{x}_j e tomando as

estimativas de máxima verossimilhança \hat{L} e $\hat{\Psi}$ como verdadeiros valores, o j -ésimo escore fatorial é dado por

$$\hat{f}_{\sim j} = \hat{L}' \hat{\Sigma}^{-1} (\mathbf{x}_{\sim j} - \bar{\mathbf{x}}) = \hat{L}' \left(\hat{L} \hat{L}' + \hat{\Psi} \right)^{-1} (\mathbf{x}_{\sim j} - \bar{\mathbf{x}}), \quad j = 1, 2, \dots, n \quad (\text{III})$$

De acordo com o resultado 2.30, o cálculo de $\hat{f}_{\sim j}$ em (III) pode ser simplificado pelo uso da matriz identidade:

$$\hat{L}' \left(\hat{L} \hat{L}' + \hat{\Psi} \right)^{-1} = \left(I + \hat{L}' \hat{\Psi}^{-1} \hat{L} \right)^{-1} \hat{L}' \hat{\Psi}^{-1} \quad (\text{IV})$$

(mxp) (pxp) (mxm) (mxp) (pxp)

Esta identidade ajuda a comparar os escores fatoriais em (III), gerados por regressão, com aqueles gerados pelo Método dos Mínimos Quadrados Ponderados. Assim, denotando $\hat{f}_{\sim j}^R$ como o escore fatorial pelo Método da Regressão, e $\hat{f}_{\sim j}^{LS}$ como o escore fatorial pelo

Método dos Mínimos Quadrados Ponderados, por (IV) tem-se:

$$\hat{f}_{\sim j}^{LS} = \left(\hat{L}' \hat{\Psi}^{-1} \hat{L} \right)^{-1} \left(I + \hat{L}' \hat{\Psi}^{-1} \hat{L} \right) \hat{f}_{\sim j}^R = \left(I + \left(\hat{L}' \hat{\Psi}^{-1} \hat{L} \right)^{-1} \right) \hat{f}_{\sim j}^R \quad (\text{V})$$

Para as estimativas de máxima verossimilhança, $\left(\hat{L}' \hat{\Psi}^{-1} \hat{L} \right)^{-1} = \hat{\Delta}^{-1}$, e se os elementos desta matriz diagonal estão próximos de zero, o Método da Regressão e o de Mínimos Quadrados Generalizados resultarão em escores fatoriais muito próximos.

Na tentativa de reduzir os possíveis efeitos da incorreta determinação do número de fatores, práticas tendem a calcular os escores fatoriais de (V) pelo uso de S , a matriz de covariância amostral original, em vez de $\hat{\Sigma} = \hat{L} \hat{L}' + \hat{\Psi}$. Pode-se, assim, montar o Modelo para o cálculo dos escores fatoriais por regressão da seguinte maneira.

$$\bullet \hat{f}_{\sim j} = L' S^{-1} (\underline{x}_{\sim j} - \bar{\underline{x}}), \quad j = 1, 2, \dots, n$$

ou, se a matriz de correlação é fatorada,

$$\bullet \hat{f}_{\sim j} = L_z' R^{-1} z_{\sim j}, \quad j = 1, 2, \dots, n$$

onde

$$z_{\sim j} = D^{-1/2} (\underline{x}_{\sim j} - \bar{\underline{x}}) \quad \text{e} \quad \hat{\rho} = L_z L_z' + \hat{\Psi}_z$$

Caso os carregamentos rotacionados $\hat{L}^* = \hat{L}T$ sejam usados no lugar dos carregamentos originais, os subseqüentes escores fatoriais $\hat{f}_{\sim j}^*$ são relacionados a $\hat{f}_{\sim j}$ por

$$\hat{f}_{\sim j}^* = T \hat{f}_{\sim j}, \quad j = 1, 2, \dots, n$$

Uma medida numérica em concordância entre os escores fatoriais dos dois diferentes métodos de cálculo é produzida pelo coeficiente de correlação amostral entre escores do mesmo fator. Dos métodos apresentados até aqui, nenhum deles é recomendado como uniformemente superior.

Resultado 2.25:

Se \underline{X} é um vetor aleatório, \underline{c} é um vetor de constantes, Σ é a matriz de covariância de \underline{X} e as componentes de \underline{X} são variáveis aleatórias com média μ_i , independentes e identicamente distribuídas, então

$$V(\underline{c}' \underline{X}) = \underline{c}' \Sigma \underline{c}$$

Prova:

Seja \tilde{X}_1 um vetor aleatório e c uma constante real, das propriedades da esperança e da variância, pode-se escrever:

$$E(c \tilde{X}_1) = c E(\tilde{X}_1) = c \tilde{\mu}_1$$

e

$$V(c \tilde{X}_1) = E(c \tilde{X}_1 - c \tilde{\mu}_1)^2 = c^2 V(\tilde{X}_1) = c^2 \Sigma$$

Se \tilde{X}_2 é um outro vetor aleatório e c_1 e c_2 são constantes reais, então, das propriedades da covariância, tem-se:

$$\text{Cov}(c_1 \tilde{X}_1, c_2 \tilde{X}_2) = E(c_1 \tilde{X}_1 - c_1 \tilde{\mu}_1)(c_2 \tilde{X}_2 - c_2 \tilde{\mu}_2) = c_1 c_2 \text{Cov}(\tilde{X}_1, \tilde{X}_2)$$

Se $c' = [c_1, c_2]$, então a combinação linear $c_1 \tilde{X}_1 + c_2 \tilde{X}_2$ pode ser escrita da seguinte forma

$$[c_1 \quad c_2] \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix} = c' \tilde{X}$$

Da mesma maneira, a esperança da combinação linear seria

$$E(c_1 \tilde{X}_1 + c_2 \tilde{X}_2) = c_1 \tilde{\mu}_1 + c_2 \tilde{\mu}_2 = [c_1 \quad c_2] \begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{bmatrix} = c' \tilde{\mu}$$

Assim

$$V(c_1 \tilde{X}_1 + c_2 \tilde{X}_2) = V(c' \tilde{X}) = c' \text{Cov}(\tilde{X}) c = c' \Sigma c$$

Resultado 2.26:

Se \underline{X} é um vetor aleatório de variáveis independentes e identicamente distribuídas e \underline{c} é um vetor de constantes, então

$$E(\underline{c}' \cdot \underline{X}) = \underline{c}' \cdot \underline{\mu}$$

onde o vetor $\underline{\mu}$ é a $E(\underline{X})$.

Prova:

Seja um vetor de variáveis aleatórias $\underline{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$ e seja um vetor de constantes

$\underline{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$. Então, utilizando propriedades da esperança, tem-se :

$$E(\underline{c}' \cdot \underline{X}) = E(c_1x_1 + c_2x_2 + \dots + c_px_p) = E(c_1x_1) + E(c_2x_2) + \dots + E(c_px_p)$$

$$E(\underline{c}' \cdot \underline{X}) = c_1E(x_1) + c_2E(x_2) + \dots + c_pE(x_p) = (c_1, c_2, \dots, c_p) \cdot \begin{bmatrix} E(x_1) \\ E(x_2) \\ \vdots \\ E(x_p) \end{bmatrix}$$

$$E(\underline{c}' \cdot \underline{X}) = (c_1, c_2, \dots, c_p) \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

Logo

$$E(\underline{c}' \underline{X}) = \underline{c}' E(\underline{X}) = \underline{c}' \underline{\mu}$$

Resultado 2.27:

Se \underline{X} tem distribuição $N_p(\underline{\mu}, \Sigma)$, as q combinações lineares

$$\underset{(q \times p)}{A} \underset{(p \times 1)}{\underline{X}} = \begin{bmatrix} a_{11}X_1 + \dots + a_{1p}X_p \\ a_{21}X_1 + \dots + a_{2p}X_p \\ \vdots \\ a_{q1}X_1 + \dots + a_{qp}X_p \end{bmatrix}$$

têm distribuição $N_q(A\underline{\mu}, A\Sigma A')$. Além disso, se \underline{d} é um vetor de constantes, então $\underset{(p \times 1)}{\underline{X}} + \underset{(p \times 1)}{\underline{d}}$

tem distribuição $N_p(\underline{\mu} + \underline{d}, \Sigma)$.

Prova:

Dos resultados 2.25 e 2.26, as combinações lineares $\underline{Z} = A\underline{X}$ têm

$$\underline{\mu}_Z = E(\underline{Z}) = E(A\underline{X}) = A\underline{\mu}_X$$

e

$$\Sigma_Z = Cov(\underline{Z}) = Cov(A\underline{X}) = A\Sigma_X A'.$$

Qualquer combinação linear $b'(AX)$ é uma combinação linear de X , da forma $a'X$ com $a = A'b$. Se X é normalmente distribuído com $N_p(\mu, \Sigma)$, então qualquer combinação linear das variáveis $a'X = a_1X_1 + a_2X_2 + \dots + a_pX_p$ tem distribuição $N(a'\mu, a'\Sigma a)$. Além disso, se $a'X$ tem distribuição $N_p(a'\mu, a'\Sigma a) \forall a \in IR^p$, então X tem distribuição $N_p(\mu, \Sigma)$. Desta forma, as conclusões de AX estão satisfeitas.

A segunda parte da prova pode ser obtida considerando $a'(X+d) = a'X + (a'd)$, onde $a'X$ tem distribuição $N(a'\mu, a'\Sigma a)$. Sabe-se do caso univariado que a adição de uma constante $a'd$ à variável $a'X$ translada a média para $a'\mu + a'd = a'(\mu + d)$, sem alterar a variância. Já que a é arbitrário, $X + d$ tem distribuição $N_p(\mu + d, \Sigma)$.

Resultado 2.28:

Seja $X = \begin{bmatrix} X_1 \\ \dots \\ X_2 \end{bmatrix}$ com distribuição $N_p(\mu, \Sigma)$ na qual $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \dots & \Sigma_{12} \\ \dots & \dots & \dots \\ \Sigma_{21} & \dots & \Sigma_{22} \end{bmatrix}$, e

$|\Sigma_{22}| > 0$. Então a distribuição condicional de X_1 , dado $X_2 = x_2$, é normal com

$$\text{Média} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$

e

$$\text{Covariância} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Assim, a covariância independe do valor de x_2 da variável condicionada.

Prova:

Seja

$$A = \begin{matrix} & \begin{matrix} I & \vdots & -\Sigma_{12}\Sigma_{22}^{-1} \\ (qxq) & & qx(p-q) \\ \dots & & \dots \\ 0 & \vdots & I \\ (p-q) \times q & & (p-q) \times (p-q) \end{matrix} \end{matrix}$$

Então

$$A(X - \mu) = A \begin{bmatrix} X_1 - \mu_1 \\ \dots \\ X_2 - \mu_2 \end{bmatrix} = \begin{bmatrix} X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2) \\ \dots \\ X_2 - \mu_2 \end{bmatrix}$$

é normal conjuntamente com matriz de covariância $A\Sigma A'$ dada por

$$\begin{bmatrix} I & \vdots & -\Sigma_{12}\Sigma_{22}^{-1} \\ \dots & & \dots \\ 0 & \vdots & I \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \vdots & \Sigma_{12} \\ \dots & & \dots \\ \Sigma_{21} & \vdots & \Sigma_{22} \end{bmatrix} \begin{bmatrix} I & \vdots & 0' \\ \dots & & \dots \\ (-\Sigma_{12}\Sigma_{22}^{-1}) & \vdots & I \end{bmatrix} = \begin{bmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & \vdots & 0' \\ \dots & & \dots \\ 0 & \vdots & \Sigma_{22} \end{bmatrix}$$

Como $X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2)$ e $X_2 - \mu_2$ tem covariância zero, são independentes.

Além disso, $X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2)$ tem distribuição $N_q(0, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$. Dado

$X_2 = \mu_2$, $\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$ é uma constante. A distribuição condicionada de

$X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$ é a mesma da distribuição não condicionada

$X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2)$, pois $X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2)$ e $X_2 - \mu_2$ são independentes. Já

que a quantidade $X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2)$ tem distribuição $N_q(0, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$, ela é o

vetor aleatório $X_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$, no qual X_2 assume o valor particular x_2 . De forma

equivalente, dado que $X_2 = x_2$, X_1 tem distribuição $N_q(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$.

Resultado 2.29:

Supondo não singulares as matrizes e possíveis as operações, então

$$(I + L'\Psi^{-1}L)^{-1} L'\Psi^{-1}L = I - (I + L'\Psi^{-1}L)^{-1}$$

Prova:

Partindo da igualdade entre matrizes identidades de mesma ordem

$$I = I$$

multiplicando membro a membro por $L'\Psi^{-1}L$, vem

$$L'\Psi^{-1}L = L'\Psi^{-1}L$$

somando e subtraindo a identidade ao segundo membro

$$L'\Psi^{-1}L = I + L'\Psi^{-1}L - I$$

$$L'\Psi^{-1}L = (I + L'\Psi^{-1}L) - (I + L'\Psi^{-1}L)(I + L'\Psi^{-1}L)^{-1}$$

fatorando a expressão :

$$L'\Psi^{-1}L = (I + L'\Psi^{-1}L) \cdot [I - (I + L'\Psi^{-1}L)^{-1}]$$

pré-multiplicando membro a membro por $(I + L'\Psi^{-1}L)^{-1}$, tem-se

$$(I + L'\Psi^{-1}L)^{-1}(L'\Psi^{-1}L) = (I + L'\Psi^{-1}L)^{-1} \cdot (I + L'\Psi^{-1}L) \cdot [I - (I + L'\Psi^{-1}L)^{-1}]$$

como $(I + L'\Psi^{-1}L)^{-1} \cdot (I + L'\Psi^{-1}L) = I$, pode-se escrever

$$(I + L'\Psi^{-1}L)^{-1}(L'\Psi^{-1}L) = I - (I + L'\Psi^{-1}L)^{-1}$$

Resultado 2.30:

Supondo não singulares as matrizes e possíveis as operações, então

$$(LL' + \Psi)^{-1} = \Psi^{-1} - \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}$$

Prova:

Partindo da igualdade entre matrizes identidades de mesma ordem

$$I = I$$

somando e subtraindo expressões ao segundo membro, tem-se

$$I = I + \Psi^{-1}LL' - \Psi^{-1}LL' + \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L' - \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L'$$

$$I = I + \Psi^{-1}LL' - \Psi^{-1}L.[I - (I + L'\Psi^{-1}L)^{-1}]L' - \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L'$$

usando o resultado 2.29 e substituindo $I = \Psi^{-1}\Psi$, tem-se

$$I = \Psi^{-1}\Psi + \Psi^{-1}LL' - \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}LL' - \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L'$$

fatorando

$$I = \Psi^{-1} (LL' + \Psi) - \Psi^{-1}L. [(I + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}LL' + (I + L'\Psi^{-1}L)^{-1}L']$$

fatorando

$$I = \Psi^{-1} (LL' + \Psi) - \Psi^{-1}L. [(I + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}](LL' + \Psi)$$

pós-multiplicando membro a membro por $(LL' + \Psi)^{-1}$, tem-se

$$(LL' + \Psi)^{-1} = \Psi^{-1} - \Psi^{-1}L.(I + L'\Psi^{-1}L)^{-1}L'\Psi^{-1}$$

2.4.8.3 Método da Regressão para Fatores Correlacionados

O Método da Regressão na estimação de f , vetor de escores fatoriais, pode, sem dificuldade, ser generalizado a casos nos quais os fatores são correlacionados e não necessariamente padronizados.

Denotando a matriz de covariância entre os fatores por Φ , tem-se

$$E(\underline{\hat{f}} \underline{\hat{f}}') = \Phi$$

$$E(\underline{x} \underline{\hat{f}}') = E[(L \underline{f} + \underline{\varepsilon}) \underline{\hat{f}}'] = L \Phi$$

$$E(\underline{x} \underline{x}') = \Sigma = L \Phi L' + \Psi$$

Para $m = 1, 2, \dots, k$, procura-se uma função linear das observações que produza um satisfatório valor esperado $\underline{\hat{f}}_m$. Seja este valor

$$\underline{\hat{f}}_m = \underline{a}_m' \underline{x} = \underline{x}' \underline{a}_m$$

no qual \underline{a}_m é um vetor de ordem p . Deve-se escolher \underline{a}_m de forma que $\underline{\hat{f}}_m - \underline{f}_m \Phi$ seja minimizado. Esta variância é

$$E\left(\underline{\hat{f}}_m - \underline{f}_m \Phi\right)^2 = E\left(\underline{x}' \underline{a}_m - \underline{f}_m \Phi\right)^2$$

Para minimizar a expressão anterior, iguala-se a zero sua derivada em relação ao vetor \underline{a}_m . A derivada é

$$E[2 \underline{x} (\underline{x}' \underline{a}_m - \underline{f}_m \Phi)] = 2(\Sigma \underline{a}_m - \underline{\ell}_m \Phi)$$

onde $\underline{\ell}_m$ é a m -ésima coluna de L . Assim, tem-se

$$\Sigma \underline{a}_m = \underline{\ell}_m \Phi$$

ou

$$\underline{a}_m = \Sigma^{-1} \underline{\ell}_m \Phi$$

e $\underline{\hat{f}}_m$ é dado por

$$\hat{\underset{\sim}{f}}_m = \Phi \underset{\sim}{\ell}'_m \Sigma^{-1} \underset{\sim}{x}$$

se $\hat{\underset{\sim}{f}}$ denota o vetor de estimativas $\hat{\underset{\sim}{f}}_1, \hat{\underset{\sim}{f}}_2, \dots, \hat{\underset{\sim}{f}}_k$, pode-se escrever

$$\hat{\underset{\sim}{f}} = \Phi L' \Sigma^{-1} \underset{\sim}{x}$$

ou

$$\hat{\underset{\sim}{f}} = \Phi (I + \Gamma \Phi)^{-1} L' \Psi^{-1} \underset{\sim}{x}$$

onde $\Gamma = L' \Psi^{-1} L$.

Com tudo isso, pode-se escrever

$$E\left(\hat{\underset{\sim}{f}} \hat{\underset{\sim}{f}}'\right) = E\left(\hat{\underset{\sim}{f}} \hat{\underset{\sim}{f}}\right) = \Phi L' \Sigma^{-1} L \Phi$$

$$E\left(\hat{\underset{\sim}{f}} \hat{\underset{\sim}{f}}'\right) = E\left(\hat{\underset{\sim}{f}} \hat{\underset{\sim}{f}}'\right) = \Phi - \Phi (I + \Gamma \Phi)^{-1}$$

$$E\left[\left(\hat{\underset{\sim}{f}} - \underset{\sim}{f}\right)\left(\hat{\underset{\sim}{f}} - \underset{\sim}{f}\right)'\right] = \Phi (I + \Gamma \Phi)^{-1}$$

e

$$E(\hat{\underset{\sim}{f}} | \underset{\sim}{f}) = \Phi (L' \Sigma^{-1} L) \underset{\sim}{f}$$

$$E(\hat{\underset{\sim}{f}} | \underset{\sim}{f}) = \Phi (I + \Gamma \Phi)^{-1} (L' \Psi^{-1} L) \underset{\sim}{f}$$

$$E(\hat{\underset{\sim}{f}} | \underset{\sim}{f}) = (I + \Phi \Gamma)^{-1} \Phi \Gamma \underset{\sim}{f}$$

$$E(\hat{\underset{\sim}{f}} | \underset{\sim}{f}) = \underset{\sim}{f} - (I + \Phi \Gamma)^{-1} \underset{\sim}{f}$$

2.4.9 PERSPECTIVAS E UMA ESTRATÉGIA PARA A ANÁLISE FATORIAL

Existem muitas decisões a serem tomadas num estudo analítico de fatores. Segundo JOHNSON & WICHERN (1992), provavelmente a decisão mais importante é a escolha do número de fatores m . Mesmo que o teste realizado em uma amostra grande esteja disponível para um dado número de fatores m , isto só é possível para dados que estejam aproximadamente normalmente distribuídos. Além do mais, o teste seria rejeitado para um número muito grande de variáveis e observações. Assim sendo, a escolha final do número m está baseada numa combinação de:

- a) proporção da variância amostral explicada;
- b) conhecimento do assunto ao qual se faz a análise;
- c) significância dos resultados.

A escolha do método e tipo de rotação são decisões menos cruciais. Na prática, a maioria das análises fatoriais que apresentam um bom ajuste dos dados são aquelas nas quais as rotações são aplicadas a mais de um método, e todos os resultados substancialmente confirmam a mesma estrutura fatorial.

Pela complexidade com que se estabelece a estrutura fatorial, considerando todos os parâmetros envolvidos, dispondo de diferentes métodos, analisando significativamente os resultados, e, ainda, levando em conta a subjetividade das interpretações, pode-se adotar um procedimento estratégico para uma melhor análise da estrutura de dados.

- 1) Performance da Análise Fatorial de Componentes Principais. Este método é particularmente apropriado como primeiro passo na análise dos dados. Principalmente porque não requer que a matriz de correlação R ou a matriz de covariância S sejam não singulares:

- a) em seguida, deve-se examinar as observações mais importantes e criar gráficos a fim de relacionar os escores fatoriais. Além disso, deve-se também calcular os escores padronizados fatoriais para cada observação, estabelecendo o relacionamento entre as distâncias quadradas;
 - b) tentar uma Rotação Varimax.
- 2) O próximo passo é verificar a performance da Análise Fatorial de Máxima Verossimilhança, incluindo a Rotação Varimax.
 - 3) Comparar as soluções da Análise Fatorial em cada método utilizado:
 - a) verificar se os grupos de carregamentos representam a mesma estrutura;
 - b) localizar em gráficos os escores fatoriais obtidos por Componentes Principais pela Análise de Máxima Verossimilhança.
 - 4) Repetir os três primeiros passos para outro número de fatores comuns. Verificar se os fatores extras observados, necessariamente contribuem para um melhor entendimento e interpretação dos dados.
 - 5) Para conjuntos de dados maiores, pode-se dividir ao meio e observar a performance da Análise Fatorial de cada parte. Comparar entre si as duas soluções e então obter dos dados completos uma checagem da estabilidade da solução. Esses dados podem ser divididos aleatoriamente ou pela metade, considerando a primeira metade em um grupo, e a outra em um segundo grupo.

3 MATERIAIS E MÉTODOS

3.1 COLETA E PREPARAÇÃO DOS DADOS

Os métodos abordados no capítulo 2 foram aplicados a dados de um tipo de Área Especialmente Protegida: Faxinais do Estado do Paraná. Os dados foram obtidos por DOMINGUES (1999), coletados de 20 Faxinais pertencentes a 4 municípios, distribuídos conforme o Quadro 3.

QUADRO 3 - LISTAGEM DOS FAXINAIS ESTUDADOS POR MUNICÍPIO

Município	Faxinais
Prudentópolis	Linha Ivai-Anta Gorda Papanduva De Baixo Cachoeira do Palmital Tijuco Preto Taboãozinho Marcondes (Linha Brasília) Linha Paraná-Anta Gorda Guanabara Rio Bonito Rio do Meio São Pedro Ponte Nova Patos Velhos Queimados
Rebouças	Marmeleiro de Cima Marmeleiro de Baixo Salto
Irati	Rio dos Couros Dos Mellos
Boa Ventura de São Roque	Dos Krieger

Uma vez que o Decreto Estadual N.º3.446/97 foi instituído apenas no ano de 1997, a existência de Áreas Especiais de Uso Regulamentado – ARESUR passaram a valer, para fins de ICMS Ecológico, apenas em 1998. A partir de 1999, o Instituto Ambiental do Paraná – IAP passou a utilizar o critério qualitativo para a determinação dos escores de cada ARESUR.

3.1.1 AMOSTRAGEM

O universo considerado foi o total de Faxinais de todos os municípios estudados. O grau de confiança estimado foi de 90%; e a precisão da estimativa de 3%.

O número de famílias a serem entrevistadas em todo o universo considerado foi dimensionado partindo-se da estimação do parâmetro de uma variável aleatória Bernoulli, definido por:

$$n = \frac{0,25N}{0,25 + \frac{d}{Z} \cdot (N - 1)}$$

no qual:

- n é o número de amostras a serem coletadas do universo;
- N é o tamanho da população;
- d é a precisão da estimativa;
- Z é o escore da curva normal padrão correspondendo ao nível de confiança.

O número de famílias a serem entrevistadas em cada Faxinal foi definido, de forma proporcional, de acordo com a seguinte formulação:

$$n_i = \frac{n \cdot N_i}{N}$$

no qual:

- n_i é o número de amostras a serem coletadas no Faxinal i ;
- n é o número de amostras a serem coletadas do universo;
- N_i é o tamanho da população do Faxinal i ;
- N é o tamanho da população composta por todos os Faxinais (universo).

3.1.2 ELABORAÇÃO DO QUESTIONÁRIO

A elaboração do questionário para a coleta dos dados foi realizada por DOMINGUES (1999). Fundamentalmente foram aplicados dois questionários. Sendo um deles as unidades familiares; e outro, as lideranças locais. O conteúdo do questionário foi baseado, segundo a mesma autora, em variáveis definidas para se atingirem os objetivos de criar critérios para um ranqueamento dos Faxinais, de acordo com o IAP. Os questionários encontram-se anexos ao trabalho.

3.1.3 DEFINIÇÃO DAS VARIÁVEIS

O critério qualitativo adotado pelo IAP a partir de 1999 depende de variáveis presentes nos Faxinais, tais como densidade populacional, qualidade de vida das populações residentes, organização comunitária e nível de comprometimento e empenho dos municípios para com o desenvolvimento social e econômico dos Faxinais.

As variáveis no presente estudo foram sugeridas por DOMINGUES (1999), considerando os conceitos da literatura e o Decreto Estadual 3.446/97.

As variáveis consideradas são as seguintes:

1. **proprietários** – número de produtores do Faxinal que possuem a propriedade da terra, em percentual;
2. **arrendatários, parceiros ou agregados** – número de agricultores que se declararam somente arrendatários, parceiros ou agregados, em percentual;
3. **arrenda ou parceria** – número de produtores do Faxinal que, mesmo sendo proprietários, utilizam-se da terra de terceiros para a produção agrícola, através de arrendamento ou parceria, em percentual;
4. **renda média** – renda bruta anual média, em reais;
5. **renda da agricultura** – renda bruta anual média proveniente da agricultura, em reais;
6. **renda de outras fontes** – renda bruta anual média proveniente de outras fontes que não sejam a agricultura, salários, aposentadorias ou pensões e diárias, em reais;
7. **áreas de lavoura** – médias de áreas utilizadas para lavoura, em hectare;
8. **áreas de floresta nativa** – média das áreas de florestas nativas, em hectare;
9. **áreas de floresta plantada** – média das áreas de florestas plantadas, em hectare;
10. **áreas em descanso** – média das áreas em descanso, em hectare;
11. **áreas no criadouro comunitário** – médias das áreas utilizadas no criadouro comunitário, em hectares;
12. **consumo de leite** – consumo diário de leite, em percentual;

13. **consumo de derivados de leite** – consumo diário de derivados de leite, como manteiga e queijos, em percentual;
14. **consumo de ovos** – consumo diário de ovos, em percentual;
15. **consumo de carne de frango** – consumo diário de carne de frango, em percentual;
16. **consumo de carne de gado** – consumo diário de carne de gado, em percentual;
17. **consumo de carne de porco** - consumo diário de carne de porco, em percentual;
18. **consumo de batata/mandioca** – consumo diário de batata, mandioca e similares, em percentual;
19. **consumo de farinha de trigo** – consumo diário de farinha de trigo e derivados, em percentual;
20. **consumo de milho e derivados** – consumo diário de milho e derivados, em percentual;
21. **consumo de hortaliças** – consumo diário de hortaliças de folha, em percentual;
22. **consumo de legumes** – consumo diário de legumes, em percentual;
23. **consumo de frutas** – consumo diário de frutas, em percentual;
24. **consumo de doces** – consumo diário de doces, bolos e derivados, em percentual;
25. **consumo de óleo** – consumo de óleo vegetal, em percentual;
26. **produção de erva-mate** – produção média de erva mate, em arroba;
27. **produtor de erva mate** – número de produtores de erva mate, em percentual;
28. **produtor de bracatinga** – número de produtores de lenha de bracatinga, em percentual;
29. **consumo de lenha** – quantidade média de lenha consumida, em st;
30. **produtor de eucalipto para lenha** – número de agricultores que plantaram eucalipto para lenha, em percentual;

31. **plantam araucária** – número de agricultores que plantaram *Araucaria angustifolia*, em percentual;
32. **mudas de araucária plantadas** – quantidade média de mudas de *Araucaria angustifolia* plantada;
33. **plantam erva mate** – número de agricultores que plantaram erva mate, em percentual;
34. **mudas de erva mate plantadas** – quantidade média de mudas de erva mate plantada;
35. **plantam exóticas** – número de agricultores que plantaram mudas de espécies exóticas, em percentual;
36. **mudas de exóticas plantadas** – quantidade média de mudas de espécies exóticas plantadas;
37. **conservação de solos** – quantidade de produtores que utilizaram alguma prática de conservação de solos, em percentual;
38. **calagem de solo** – quantidade de produtores que fizeram calagem de solos, em percentual;
39. **sementes selecionadas** – quantidade de produtores que utilizaram sementes selecionadas, em percentual;
40. **crédito agrícola** – produtores que utilizaram crédito agrícola, em percentual;
41. **assistência técnica** – produtores que receberam algum tipo de assistência técnica, em percentual;
42. **renda de milho** – renda bruta anual média obtida com o milho, em reais;
43. **renda de feijão** – renda bruta anual média obtida com feijão, em reais;
44. **renda do fumo** – renda bruta anual média obtida com o fumo, em reais;

45. **renda da erva mate** – renda bruta anual média obtida com a erva mate, em reais;
46. **número de bovinos** – número de cabeças de bovinos, média;
47. **número de animais de trabalho** – número de cabeças de animais de trabalho, média;
48. **número de suínos** – número de cabeças de suínos, média;
49. **número de galinhas** – número de cabeças de galinhas, galos, e outras aves, média;
50. **filiado à associação** – quantidade do produtor filiado à associação comunitária, em percentual;
51. **participação em atividades comunitárias** – número de produtores que participam de atividades comunitárias, em percentual;
52. **dias gastos em atividades comunitárias** – média de dias do mês que o produtor gastou em atividades comunitárias;
53. **distância mínima da escola** – distância mínima da residência do produtor a uma escola de 1ª a 4ª série, em Km, que foi modificada para atender os objetivos da análise, da seguinte forma: 0 a 1,99 Km =3; 2 a 3,99Km=2; 4 a 6Km=1;
54. **distância mínima do posto de saúde** – distância mínima da residência do produtor a um posto de saúde, em Km, também modificada para atender os objetivos da análise, da seguinte forma: 0 a 8,99Km=3; 9 a 17,99Km=2; 18 a 27 Km=1;
55. **distância mínima do asfalto** – distância mínima, através de estradas, da residência do produtor à estrada de asfalto mais próxima, em Km, também modificada, da seguinte forma: 0 a 5,99Km=3; 6 a 11,99Km=2; 12 a 18Km=1;
56. **abastecimento de água com bomba elétrica** – número de residências, cujo abastecimento de água é feito através de poço com bomba elétrica, em percentual;

57. **rede pública comunitária de água** – número de residências, cujo abastecimento de água é feito através de rede pública comunitária, em percentual;
58. **luz de rede pública** – número de residências que possuem banheiro interno, em percentual;
59. **banheiro interno** – número de residências que possuem luz elétrica de rede pública, em percentual;
60. **esgoto com fossa séptica** – número de residências que possuem fossa séptica para esgotamento sanitário, em percentual;
61. **sindicato para assistência médica** – número de famílias que utilizam do sindicato para assistência médica, em percentual;
62. **rede pública para assistência médica** – número de famílias que se utilizam da rede pública para assistência médica, em percentual;
63. **tamanho da casa** – tamanho médio das residências, em m² ;
64. **sempre moraram na comunidade** – número de produtores que sempre moraram na comunidade, em percentual;
65. **uso múltiplo** – número de produtores que mencionam mais de uma forma de uso para a floresta, em percentual;
66. **desenvolver e proteger** – número de produtores que acham possível promover o desenvolvimento e proteger o meio ambiente ao mesmo tempo, em percentual;
67. **futuro dos filhos no campo** – número de produtores que acham que o futuro dos filhos será no campo, em percentual;
68. **nunca pensaram em deixar de ser agricultor** – número de agricultores que declararam nunca ter pensado em deixar de ser a vida no campo, em percentual;

69. **nunca pensaram em acabar com o Faxinal** – número de agricultores que declararam nunca ter pensado em acabar com o Faxinal onde moram, em percentual;
70. **situação piora se o faxinal acabar** – número de agricultores que acreditam que sua situação vai piorar se o Faxinal acabar, em percentual;
71. **criação solta como condição para faxinal** – produtores que pensam que para ser faxinal tem que ter criação solta, em percentual;
72. **produtores alfabetizados** – número de produtores alfabetizados, em percentual;
73. **melhoramento genético da criação** – se no Faxinal ocorreu melhoria genética da criação nos últimos 5 anos, atribuindo valor “1” se ocorreu, e valor “0” se não ocorreu;
74. **possui criadouro comunitário** – se o Faxinal possui criadouro comunitário ativo completo, atribuindo o valor “1” se possui, e valor “0” se não possui;
75. **conhecimento do ICMS Ecológico** – se a comunidade está ciente e participou da elaboração do plano de ação para o ICMS Ecológico, conforme prevê o Decreto Estadual no. 3.446/97, atribuindo valor “1” se conhecem, e valor “0” se não conhecem;
76. **possui associação** – se o Faxinal possui associação comunitária, atribuindo valor “1” se possui, e valor “0” se não possui;
77. **possui escola** – se o Faxinal possui escola, atribuindo valor “0” quando não possui, valor “1” quando a escola é até a 4ª série e valor “2” quando a escola é até a 8ª série;
78. **possui posto de saúde** – se o Faxinal possui posto de saúde ativo, atribuindo valor “1” se possui, e valor “0” se não possui;

79. **densidade** – considerou-se a densidade populacional do criadouro, conforme prevê o Decreto Estadual 3.446/97, ou seja, o número de moradores dividido pela área, em hectare, do criadouro comunitário;
80. **possui equipamentos comunitários** – se a comunidade do Faxinal possui equipamentos de uso comunitários como caminhões, máquinas agrícolas, cozinhas, viveiros, entre outros, atribuindo valor “1” se possui, e valor “0” se não possui.

3.2 MÉTODO ESTATÍSTICO DE ANÁLISE

Os dados que foram coletados referem-se ao ano agrícola de agosto de 1997 a julho de 1998. Os questionários propostos e as variáveis definidas foram sugeridas a partir de um estudo feito por DOMINGUES (1999).

Inicialmente os dados são coletados através de questionários aplicados às famílias e lideranças locais. Foram aplicados 316 (trezentos e dezesseis) questionários em um universo de 1947 famílias. Segundo a mesma autora, a precisão da estimativa relativa ao universo considerado foi igual a 0,055 levando-se em conta um grau de confiança de 95%, o que fez com que se obtivesse a intensidade de acordo com a Quadro 4.

QUADRO 4 - LISTAGEM DO NÚMERO DE FAMÍLIAS, NÚMERO DE AMOSTRAS NECESSÁRIAS E TOMADAS

Faxinais	N.º de famílias	N.º de amostras necessárias (n_h)	N.º de amostras tomadas
Marmeleiro de Baixo	135	19	19
Marmeleiro de Cima	48	7	7
Salto	50	7	7
Dos Krieger	45	6	10
Dos Mellos	32	5	5
Rio dos Couros	65	9	9
Ponte Nova	50	7	7
Guanabara	45	6	8
Queimados	150	21	23
Papanduva de Baixo	135	19	19
Marcondes (Linha Brasília)	80	11	12
Taboãozinho	78	11	12
Patos Velhos	87	12	12
Linha Paraná-Anta Gorda	45	6	22
Cachoeira do Palmital	140	20	25
Rio Bonito	8	1	3
São Pedro	85	12	12
Rio do Meio	65	9	9
Linha Ivaí-Anta Gorda	54	8	18
Tijuco Preto	550	77	77
Total	1947	273	316

Para o ranqueamento dos Faxinais, foi construído um programa computacional que estima os escores de cada Faxinal através da Análise Fatorial, abordada na seção 2.2.

A partir dos dados coletados, foi criada uma matriz de dados de ordem 20x80, sendo 20 o número de Faxinais considerados, e 80 o número de variáveis. Dessa matriz, calcularam-se as correlações existentes entre cada par de variáveis, criando, assim, a matriz de correlações. Em seguida, determinaram-se os pares de autovalores e autovetores dessa matriz de correlação. Adotou-se o critério 2, da página 45 do presente estudo, ou seja, o número de fatores considerado foi igual ao número de autovalores maiores que 1 da matriz de correlação.

Desta forma, foram considerados 17 fatores, o que representou um grau de explicação de 97,7649%.

A matriz de carregamentos fatoriais foi formada a partir dos autovalores e autovetores associados. Os fatores foram, então, sem alteração do grau de explicação, rotacionados pelo Método Varimax Normal, sendo obtida, assim, a matriz de carregamentos fatoriais rotacionados. Em seguida, determinaram-se as comunalidades e as variâncias específicas de cada variável.

Os escores fatoriais, que são estimativas dos Faxinais para os fatores não observáveis, foram calculados pelo Método de Mínimos Quadrados Ponderados. Os escores de cada Faxinal foram obtidos ponderados pela importância de cada fator, representada pelo autovalor.

A escala de escores fatoriais foi alterada para atender às exigências do IAP, que determina, pela Portaria n.º 263/98/IAP/GP de 1998, que os escores de qualidade mínimo e máximo para Faxinais devem ficar entre “0” e “2”. Após a padronização, foram encontrados os escores finais.

3.3 PROGRAMA

O programa foi elaborado visando classificar qualquer tipo de categoria, considerando ainda, qualquer número de variáveis. Sendo assim, é completamente genérico e atende perfeitamente o usuário que procura ranquear, utilizando a Análise Fatorial.

O objetivo maior foi o de medir a qualidade de vida, a participação e organização comunitária e o nível de comprometimento dos municípios para com o desenvolvimento social e econômico dos Faxinais.

3.3.1 DESCRIÇÃO

O programa foi elaborado em linguagem de programação Visual Basic e em ambiente *Windows 95*.

A janela principal apresenta as opções de arquivo, cálculo, configurações do programa, janelas abertas e função de ajuda.

Uma vez que os dados originais são inseridos, podem ser obtidas as correlações entre cada par de variáveis, as comunalidades e as variâncias específicas de cada variável; os autovalores e autovetores da matriz de correlação, a matriz de carregamentos fatoriais, a matriz de carregamentos fatoriais rotacionados, os escores fatoriais, os escores brutos e os escores padronizados. Selecionando a opção *Mostrar desenvolvimento*, além dos escores, serão mostrados todos os tópicos mencionados.

Os escores podem ser armazenados acessando a opção *Armazenar Escores*. Os escores brutos e padronizados referentes a anos anteriores também podem ser requeridos. Neste caso, basta acessar a opção *Escores de Anos Anteriores*. Com isso, pode-se monitorar com grande eficiência o desenvolvimento dos escores ao longo dos anos.

Existe, ainda, uma opção para determinar o intervalo que se deseja situar o escore de cada Unidade de Conservação. Desta forma, os escores podem ser padronizados num intervalo definido pelo usuário.

Uma opção disponível no programa é a *Ajuda sobre o Programa de Ranqueamento*. Com ela o usuário pode eliminar qualquer dúvida que possa eventualmente aparecer, tornando simples e rápido o uso do programa. Além disso, é facultado ao usuário o uso da *Dica do Dia*. A *Dica do Dia* traz sempre uma mensagem que procura orientar sobre algum tópico importante ou alguma curiosidade.

O programa ainda apresenta uma tábua de avaliação que informa a identificação, a qualificação, o responsável técnico e a localização da unidade de conservação. São itens importantes numa avaliação de campo. As Figuras 6, 7, 8 e 9 mostram, respectivamente, as interfaces citadas da tábua de avaliação do programa.

Tábua de Avaliação

ICMS ECOLÓGICO
POR UNIDADES DE CONSERVAÇÃO

TÁBUA DE AVALIAÇÃO

1.3 QUALIFICAÇÃO DO FÁXINAL
1.1 IDENTIFICAÇÃO

1.4 RESPONSABILIDADE TÉCNICA DO AVALIADOR
1.2 LOCALIZAÇÃO DA UNIDADE DE CONSERVAÇÃO

Nome da Unidade de Conservação: _____
Categoria de Manejo: _____
Âmbito de Governo: _____
Data de criação da ARESUR: _____
Local: _____
Prefeito: _____
Observações: _____

Tábua de Avaliação Padrão Tábua de Avaliação Específica

OK Cancelar

FIGURA 6 - INTERFACE DE IDENTIFICAÇÃO

Tábua de Avaliação

ICMS ECOLÓGICO
POR UNIDADES DE CONSERVAÇÃO

TÁBUA DE AVALIAÇÃO

1.3 QUALIFICAÇÃO DO FAXINAL
1.1 IDENTIFICAÇÃO

1.4 RESPONSABILIDADE TÉCNICA DO AVALIADOR
1.2 LOCALIZAÇÃO DA UNIDADE DE CONSERVAÇÃO

Nome do Município de interface:

Área total do município de interface:

OK Cancelar

FIGURA 7 - INTERFACE DE LOCALIZAÇÃO DA UNIDADE DE CONSERVAÇÃO

Tábua de Avaliação

ICMS ECOLÓGICO
POR UNIDADES DE CONSERVAÇÃO

TÁBUA DE AVALIAÇÃO

1.1 IDENTIFICAÇÃO
1.3 QUALIFICAÇÃO DO FAXINAL

1.2 LOCALIZAÇÃO DA UNIDADE DE CONSERVAÇÃO
1.4 RESPONSABILIDADE TÉCNICA DO AVALIADOR

Área total da unidade (ha):

Área do criatório comunitário (ha):

Densidade populacional:

Nível de organização da unidade:

Nível de participação das comunidades:

Qualidade de vida:

Empenho do munic (pio):

Escore bruto:

Escore padronizado:

OK Cancelar

FIGURA 8 - INTERFACE DE QUALIFICAÇÃO DO FAXINAL

Tábua de Avaliação

ICMS ECOLÓGICO
POR UNIDADES DE CONSERVAÇÃO

TÁBUA DE AVALIAÇÃO

1.1 IDENTIFICAÇÃO | 1.2 LOCALIZAÇÃO DA UNIDADE DE CONSERVAÇÃO
1.3 QUALIFICAÇÃO DO FAXINAL | 1.4 RESPONSABILIDADE TÉCNICA DO AVALIADOR

Nome completo do profissional avaliador:

Número do registro profissional:

Escritório regional do IAP:

Local da avaliação:

Data da avaliação:

OK Cancelar

FIGURA 9 - INTERFACE DE RESPONSABILIDADE TÉCNICA DO AVALIADOR

3.3.2 FLUXOGRAMA

Existe um longo caminho computacional, desde a entrada dos dados brutos, até a obtenção dos escores padronizados finais de cada Faxinal. Para a compreensão de todos os passos efetivamente utilizados no cálculo do escore, é destacada a seguir a Figura 10, que mostra o fluxograma do cálculo do escore padronizado.

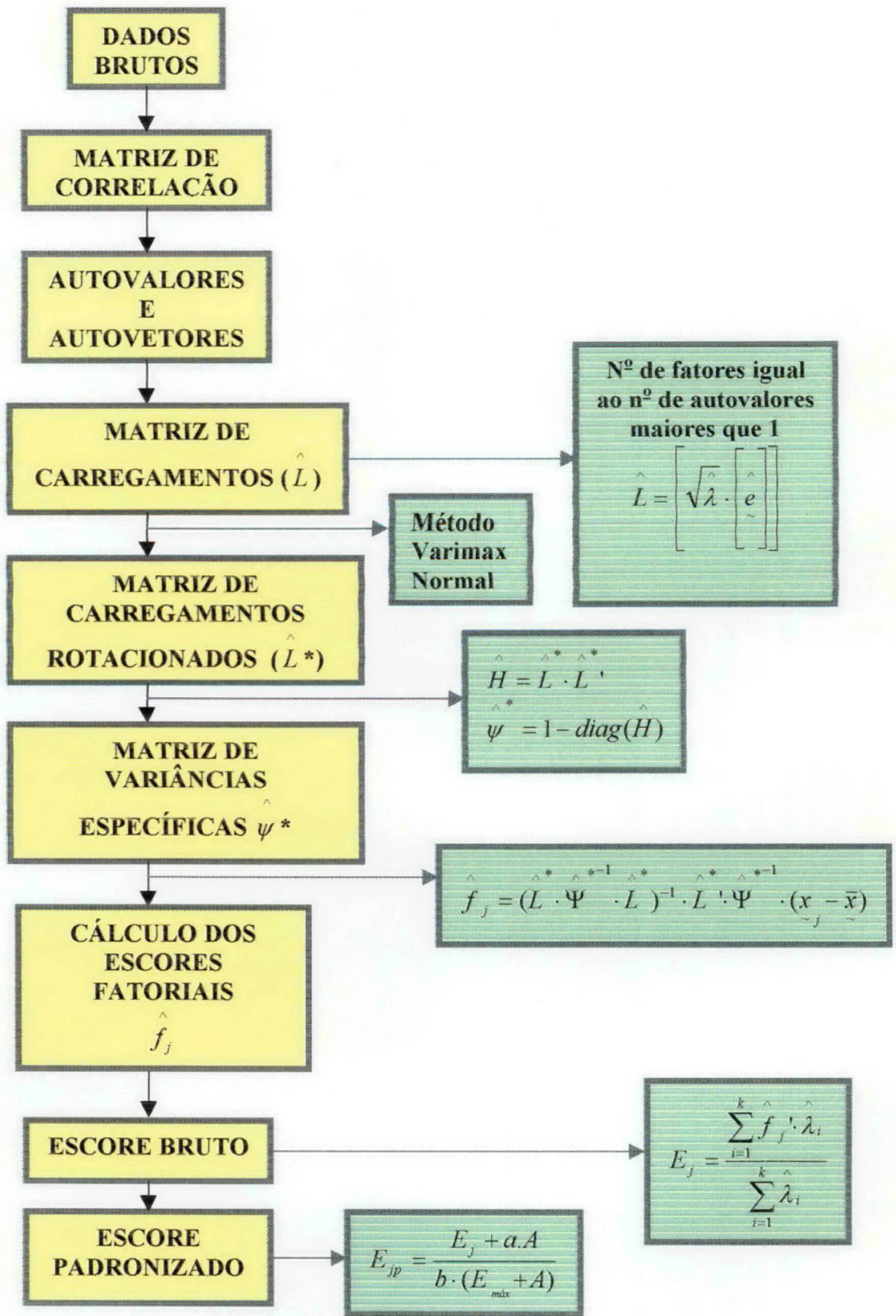


FIGURA 10 - FLUXOGRAMA DO CÁLCULO DO ESCORE PADRONIZADO

3.3.3 CÁLCULO DOS ESCORES PADRONIZADOS

Atendendo exigências do usuário, o escore padronizado E_p pode ser obtido em função do escore bruto E , da amplitude A e de dois parâmetros a e b , conforme a equação:

$$E_p = \frac{E + a.A}{b.(E_{m\acute{a}x} + A)} \quad (\text{I})$$

onde $A = E_{m\acute{a}x} - E_{m\acute{i}n}$ e $E_p = [Z_1, Z_2]$, $Z_1, Z_2 \geq 0$, $Z_1 < Z_2$.

O menor escore padronizado é calculado em função do menor escore bruto. É dado um acréscimo de 20% a Z_1 para a obtenção do menor escore mínimo.

$$E_{p_{m\acute{i}n}} = \frac{E_{m\acute{i}n} + a.A}{b.(E_{m\acute{a}x} + A)} = 1,2Z_1 \geq Z_1 \quad (\text{II})$$

O maior escore padronizado é calculado em função do maior escore bruto. É dada uma redução de 10% a Z_2 .

$$E_{p_{m\acute{a}x}} = \frac{E_{m\acute{a}x} + a.A}{b.(E_{m\acute{a}x} + A)} = 0,9Z_2 \leq Z_2 \quad (\text{III})$$

Assim, qualquer escore padronizado, seja máximo ou seja mínimo, estará compreendido entre Z_1 e Z_2 .

Resolvendo o sistema formado por (II) e (III), obtém-se os parâmetros a e b em função das extremidades do intervalo, do menor escore, do maior escore e da amplitude:

$$a = \frac{1,2Z_1 E_{m\acute{a}x} - 0,9Z_2 E_{m\acute{i}n}}{A.(0,9Z_2 - 1,2Z_1)} \quad (\text{IV})$$

$$b = \frac{A}{(0,9Z_2 - 1,2Z_1).(E_{m\acute{a}x} + A)} \quad (\text{V})$$

Substituindo (IV) e (V) em (I), tem-se:

$$Z_1 < \underbrace{\frac{0,9.Z_2(E - E_{\min}) + 1,2.Z_1(E_{\max} - E)}{A}}_{E_p} < Z_2$$

Se $Z_1 = 0$, então em (II) substitui-se $1,2Z_1$ por $0,2$. Com isso, o menor escore padronizado da Análise Fatorial passa a ser positivo.

3.3.4 MÉTODO HOUSEHOLDER

Na aplicação aos Faxinais, a estimação dos carregamentos dos fatores foi realizada pelo Método de Componentes Principais. Para a aplicação deste método, baseado na decomposição espectral, necessitava-se dos autovalores e autovetores da matriz de correlação.

Para matrizes de menor ordem, pode-se utilizar um procedimento usual baseado no cálculo de determinantes para encontrar a equação característica, solução de equações polinomiais e sistemas. Em termos computacionais, para matrizes de ordem superior, isso não é uma tarefa tão simples.

O Método de Householder reduz uma matriz simétrica de ordem $n \times n$ a uma matriz tridiagonal, também de ordem $n \times n$, utilizando $n - 2$ transformações ortogonais. Com essa redução, transformando uma matriz simétrica numa matriz tridiagonal, os autovalores e autovetores podem ser obtidos com mais facilidade. A seguir, será descrito o Método de Householder.

Inicialmente seja A uma matriz simétrica de ordem $n \times n$, o ingrediente básico do método é a matriz de Householder P , cuja forma é:

$$P = 1 - 2 \underline{w} \underline{w}'$$

no qual \underline{w} é um vetor real tal que $|\underline{w}|^2 = 1$. Esta matriz P é ortogonal, pois

$$P^2 = (1 - 2 \underline{w} \underline{w}') \cdot (1 - 2 \underline{w} \underline{w}')$$

$$P^2 = 1 - 4 \underline{w} \underline{w}' + 4 \underline{w} (\underline{w}' \underline{w}) \underline{w}'$$

Se $\underline{w} \underline{w}' = 1$, então $P^2 = 1$.

Portanto, $P = P^{-1}$. Mas $P' = P$ e, assim, $P^T = P^{-1}$, o que prova a ortogonalidade.

Pode-se reescrever P da seguinte maneira

$$P = 1 - \frac{\underline{u} \underline{u}'}{H}$$

no qual o escalar H é dado por $H \equiv \frac{1}{2} |\underline{u}|^2$ e \underline{u} pode ser um vetor qualquer.

Supondo \underline{x} um vetor composto pela primeira coluna de A , escolhe-se

$$\underline{u} = \underline{x} \mp |\underline{x}| \underline{e}_{-1}$$

no qual \underline{e}_{-1} é o vetor unitário $[1, 0, \dots, 0]$. A escolha dos sinais é feita posteriormente.

Então

$$P \cdot \underline{x} = \underline{x} - \frac{\underline{u}}{H} \cdot (\underline{x} \mp |\underline{x}| \underline{e}_{-1})'$$

$$P \cdot \underline{x} = \underline{x} - \frac{2 \underline{u} \cdot (|\underline{x}|^2 \mp |\underline{x}| \underline{x})}{2 |\underline{x}|^2 \mp 2 |\underline{x}| \underline{x}'}$$

$$P \cdot \underline{x} = \underline{x} - \underline{u}$$

$$P \cdot \underline{x} = \pm |\underline{x}| \underline{e}_{-1}$$

Isto mostra que a matriz de Householder P atua sobre um dado vetor x , cujas coordenadas são nulas, exceto a primeira.

Para reduzir uma matriz simétrica A para a forma tridiagonal, escolhe-se o vetor x na primeira matriz de Householder como sendo os $n-1$ elementos que ficam abaixo, na primeira coluna. Assim, os demais $n-2$ elementos que ficam abaixo, na primeira coluna, serão zerados:

$$P_1.A = \begin{bmatrix} 1 & \vdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & & & & \\ 0 & \vdots & & & & \\ \vdots & \vdots & & & & \\ 0 & \vdots & & & & \end{bmatrix} \begin{bmatrix} a_{11} & \vdots & a_{12} & a_{13} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{21} & \vdots & & & & \\ a_{31} & \vdots & & & & \\ \vdots & \vdots & & & & \\ a_{n1} & \vdots & & & & \end{bmatrix}$$

$$P_1.A = \begin{bmatrix} a_{11} & \vdots & a_{12} & a_{13} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ k & \vdots & & & & \\ 0 & \vdots & & & & \\ \vdots & \vdots & & & & \\ 0 & \vdots & & & & \end{bmatrix}$$

Deve-se escrever a matriz na forma dividida, com ${}^{(n-1)}P$ denotando a matriz Householder de ordem $n-1$. A quantidade k é simplesmente mais ou menos a magnitude do vetor $[a_{21}, \dots, a_{n1}]'$.

A transformação ortogonal se completa pós-multiplicando a equação pela matriz P :

$$A' = P.A.P = \begin{bmatrix} a_{11} & \vdots & k & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ k & \vdots & & & & \\ 0 & \vdots & & & & \\ \vdots & \vdots & & & & \\ 0 & \vdots & & & & \end{bmatrix}$$

Em seguida, escolhe-se para o vetor \underline{x} , na segunda matriz de Householder, os $n-2$ elementos na segunda coluna. A segunda matriz de Householder ficaria, então, da seguinte forma:

$$P_2 \equiv \begin{bmatrix} 1 & 0 & \vdots & 0 & \cdots & 0 \\ 0 & 1 & \vdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & & & \\ \vdots & \vdots & \vdots & & & \\ 0 & 0 & \vdots & & & \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ (n-2)P_2 \end{matrix}$$

A identidade representada no bloco esquerdo superior assegura que a tridiagonalização alcançada no primeiro passo não será perdida no segundo, enquanto a matriz de Householder ${}^{(n-2)}P_2$ de ordem $(n-2)$ cria uma linha e uma coluna adicionais na matriz tridiagonal de saída.

Em vez de atualizar levando em conta as multiplicações matriciais $P.A.P$, calcula-se o vetor \underline{p} , tal que:

$$\underline{p} \equiv \frac{A \cdot \underline{u}}{H}$$

então

$$A.P = A \cdot \left(1 - \frac{\underline{u} \cdot \underline{u}'}{H} \right) = A - \underline{p} \cdot \underline{u}'$$

$$A' = P.A.P = A - \underline{p} \cdot \underline{u}' - \underline{u} \cdot \underline{p}' + 2K \underline{u} \cdot \underline{u}'$$

no qual o escalar K é definido por

$$K = \frac{\underline{u}' \cdot \underline{p}}{2H}$$

Considerando

$$\underline{q} \equiv \underline{p} - \underline{K} \underline{u}$$

tem-se

$$A' = A - \underline{q} \cdot \underline{u}' - \underline{u} \cdot \underline{q}'$$

Esta é a fórmula computacionalmente usada no método.

A rotina para a redução de Householder dada a seguir inicia na n -ésima coluna de A , não na primeira, como mostrado no procedimento anterior. Detalhando um pouco mais, as equações são as seguintes : no estágio m ($m = 1, 2, \dots, n-2$), o vetor \underline{u} tem a forma

$$\underline{u}' = [a_{i1}, a_{i2}, \dots, a_{i,i-2}, a_{i,i-1} \pm \sqrt{\sigma}, 0, \dots, 0]$$

onde

$$i \equiv n - m + 1 = n, n-1, \dots, 3$$

e a quantidade σ é dada por

$$\sigma = (a_{i1})^2 + \dots + (a_{i,i-1})^2$$

Para diminuir os erros de arredondamento, escolhe-se para o sinal de σ o mesmo de $a_{i,i-1}$.

As variáveis são então computadas na seguinte ordem : $\sigma, \underline{u}, H, \underline{p}, \underline{K}, \underline{q}, A'$. Em qualquer estágio m , A é tridiagonal em suas últimas $m-1$ linhas e colunas.

3.3.5 MÉTODO QL

Uma vez que a matriz real simétrica tenha sido reduzida à forma tridiagonal, pode-se encontrar as raízes da equação polinomial característica (autovalores) $p_n(\lambda)$ diretamente.

O polinômio característico $p_n(\lambda)$ pode ser avaliado para qualquer valor λ mediante um método recursivo. Os polinômios de grau menor produzem, durante a forma recorrente, uma seqüência de Sturmian que pode ser usada para localizar os intervalos aos quais pertencem os autovalores no eixo real. A partir daí, para encontrar as raízes, o Método da Bissecção ou o Método de Newton poderiam ser utilizados para refinar os intervalos. Os correspondentes autovetores poderiam ser encontrados por uma iteração inversa.

Entretanto, para situações em que se necessita mais do que alguns autovalores e autovetores, métodos de fatorização como o QL, são considerados muito mais eficientes.

A idéia básica do algoritmo QL é que qualquer matriz real pode ser decomposta na forma

$$A = QL$$

onde Q é uma matriz ortogonal e L é uma matriz triangular inferior.

Rescrevendo os fatores na ordem oposta, tem-se:

$$A' = LQ$$

Como Q é ortogonal, a primeira equação leva a $L = Q^T A$. Assim, da equação anterior, vem:

$$A' = Q^T A Q$$

Pode-se observar que A' é uma transformação ortogonal da matriz A que preserva propriedades da matriz A tais como simetria, forma tridiagonal e forma de Hessenberg.

Desta forma, o algoritmo QL consiste na seqüência de transformações ortogonais :

$$A_s = Q_s L_s$$

e

$$A_{s+1} = L_s Q_s, \text{ pois } A_{s+1} = Q_s^T A_s Q_s$$

Para uma melhor compreensão, será descrito a seguir um teorema fundamental, justificando a consistência do método QL para uma matriz genérica A .

Teorema:

- (i) Se A tem autovalores de diferentes valores absolutos $|\lambda_i|$, então A_s tende a uma matriz triangular inferior à medida que s tende ao infinito.

Os autovalores aparecem na diagonal em ordem crescente de magnitude absoluta.

- (ii) Se A tem um autovalor $|\lambda_i|$ com multiplicidade p , então A_s tende a uma matriz triangular inferior à medida que s tende ao infinito, exceto para a matriz de um bloco diagonal de ordem p , cujos autovalores tendem a λ_i .

Para uma matriz comum, o trabalho computacional é da ordem de n^3 operações por iteração. Já para uma matriz tridiagonal, é da ordem de n operações por iteração.

Como a matriz em questão para efeito de cálculos é a matriz transformada tridiagonal, real e simétrica, todos os autovalores são reais, e se qualquer autovalor tem multiplicidade p , então deve-se ter os menores $(p-1)$ elementos zerados na sub e superdiagonal. Desta forma, a matriz pode ser dividida em submatrizes, que podem ser diagonalizadas separadamente. As dificuldades que podem surgir dos blocos diagonais num caso geral são irrelevantes.

A partir do teorema citado anteriormente, procura-se em geral um elemento da superdiagonal que convirja para zero, usando a convergência:

$$a_{ij}^{(s)} \sim \left(\frac{\lambda_i}{\lambda_j} \right)^s$$

Mesmo que $\lambda_i < \lambda_j$, a convergência pode ser lenta se λ_i está próximo de λ_j . A convergência pode ser acelerada por uma técnica de troca: se k é constante qualquer, então $A - kI$ tem autovalores $\lambda_i - k$. Decompondo pode-se escrever

$$A_s - k_s \mathbf{1} = Q_s \cdot L_s$$

de forma que

$$A_{s+1} = L_s \cdot Q_s + k_s \mathbf{1}$$

ou

$$A_{s+1} = Q_s^T \cdot A_s \cdot Q_s$$

então a convergência fica determinada pela razão

$$\frac{\lambda_i - k_s}{\lambda_j - k_s}$$

A idéia é escolher a troca k_s a cada estágio para maximizar a razão de convergência. Uma boa escolha para a primeira troca é que k_s esteja próximo do menor autovalor λ_1 . Assim, os elementos da primeira linha fora da diagonal tenderiam rapidamente a zero. Entretanto, λ_1 não é conhecido a priori. Uma estratégia recomendável na prática é calcular os autovalores da submatriz diagonal líder de A de ordem 2 e igualar k_s ao autovalor mais próximo de a_{11} .

Mais geralmente, supondo que já tenham sido encontrados $r-1$ autovalores de A , pode-se anular as $r-1$ primeiras linhas e colunas de A , deixando, assim, a matriz com seguinte aspecto:

$$A = \begin{bmatrix} 0 & & \dots & \dots & & 0 \\ & \dots & & & & \\ \vdots & & d_r & e_r & & \vdots \\ \vdots & & e_r & d_{r+1} & & \\ & & & & \dots & 0 \\ & & & & & d_{n-1} & e_{n-1} \\ 0 & & \dots & 0 & e_{n-1} & d_n \end{bmatrix}$$

Desta forma, a decomposição QL de uma matriz genérica é efetuada por uma seqüência de transformações de Householder. Para uma matriz tridiagonal, entretanto, é mais eficiente utilizar um plano de rotação P_{pq} . Usa-se a seqüência $P_{12}, P_{23}, \dots, P_{n-1,n}$ para eliminar os elementos $a_{12}, a_{23}, \dots, a_{n-1,n}$. Por simetria, os elementos subdiagonais serão também eliminados. Assim, cada Q_s é dado por um produto de planos de rotação:

$$Q_s^T = P_1^{(s)} \cdot P_2^{(s)} \cdots P_{n-1}^{(s)}$$

no qual P_i elimina $a_{i,i+1}$.

A estratégia utilizada faz com que a convergência do método seja extremamente rápida, o que o torna bastante atrativo.

3.3.6 CÓDIGOS

Além dos algoritmos, os códigos constituem uma parte extremamente importante da programação. São destacados no ANEXO 06, em linguagem de programação Visual Basic, os códigos do Método Householder, do Método QL, da Inversão de Matrizes e da Rotação Varimax proposta por Horst que foram utilizados na Análise Fatorial.

3.3.7 APLICAÇÃO AOS FAXINAIS

Na aplicação do programa aos Faxinais, de acordo com a matriz de ordem 20x80 dos dados originais (20 Faxinais e 80 variáveis) (ANEXO 03), foram encontrados os autovalores e

autovetores da matriz de correlação (ANEXO 04) e obtidos os percentuais da variância total explicada por cada um deles.

O Quadro 5 a seguir mostra os autovalores e os percentuais da variância explicada.

QUADRO 5 - AUTOVALORES MAIORES QUE 1 DA MATRIZ DE CORRELAÇÃO E PERCENTUAL DA VARIÂNCIA EXPLICADA

Fatores	Autovalores	Total da Variância Explicada (%)	Autovalores acumulados	Total da Variância Acumulada (%)
1	13,4429	16,8036	13,4429	16,8036
2	9,6624	12,0780	23,1053	28,8816
3	8,1436	10,1796	31,2489	39,0612
4	6,7664	8,4580	38,0153	47,5191
5	5,5861	6,9826	43,6014	54,5017
6	5,4038	6,7548	49,0052	61,2565
7	4,7228	5,9035	53,7280	67,1600
8	3,9624	4,9530	57,6904	72,1130
9	3,8497	4,8121	61,5401	76,9251
10	2,9851	3,7313	64,5251	80,6564
11	2,6997	3,3747	67,2249	84,0311
12	2,3245	2,9056	69,5493	86,9367
13	2,2376	2,7969	71,7869	89,7336
14	2,1382	2,6728	73,9251	92,4064
15	1,6651	2,0813	75,5902	94,4877
16	1,3706	1,7133	76,9608	96,2010
17	1,2511	1,5639	78,2119	97,7649

Os 17 fatores considerados colaboraram para uma explicação de 97,7649% do total da variância amostral. Dentre os vários critérios que poderiam ser adotados, a escolha do número de fatores baseou-se no critério 2 do item 2.4.5, ou seja, o número de fatores considerados é igual ao número de autovalores maiores que 1 da matriz de correlação amostral. Os autovalores que não foram considerados, apresentaram pequena contribuição no percentual da variância explicada.

A variância total pode ser dividida em duas partes: comunalidades (h_i^2) das variáveis e variâncias específicas (ψ_i) das variáveis. As comunalidades representam a porção da variância total que é atribuída aos fatores comuns. Já as variâncias específicas, representam a porção da variância total atribuída à aleatoriedade (erro).

O Quadro 6 a seguir ilustra as comunalidades e as variâncias específicas de cada variável.

QUADRO 6 - COMUNALIDADES E VARIÂNCIAS ESPECÍFICAS DAS VARIÁVEIS

X_i	h_i^2	Ψ_i	X_i	h_i^2	Ψ_i	X_i	h_i^2	Ψ_i
1	0,9639	0,0361	28	0,9817	0,0183	55	0,9629	0,0371
2	0,9872	0,0128	29	0,9436	0,0564	56	0,9923	0,0077
3	0,9780	0,0220	30	0,9259	0,0741	57	0,9820	0,0180
4	0,9934	0,0066	31	0,9922	0,0078	58	0,9634	0,0366
5	0,9959	0,0041	32	0,9777	0,0223	59	0,9672	0,0328
6	0,9491	0,0509	33	0,9928	0,0072	60	0,9927	0,0073
7	0,9791	0,0209	34	0,9440	0,0560	61	0,9945	0,0055
8	0,9426	0,0574	35	0,9985	0,0015	62	0,9508	0,0492
9	0,9817	0,0183	36	0,9670	0,0330	63	0,9884	0,0116
10	0,9817	0,0183	37	0,9443	0,0557	64	0,9870	0,0130
11	0,9689	0,0311	38	0,9335	0,0665	65	0,9632	0,0368
12	0,9844	0,0156	39	0,9627	0,0373	66	0,9872	0,0128
13	0,9950	0,0050	40	0,9971	0,0029	67	0,9281	0,0719
14	0,9743	0,0257	41	0,9885	0,0115	68	0,9901	0,0099
15	0,9875	0,0125	42	0,9841	0,0159	69	0,9970	0,0030
16	0,9938	0,0062	43	0,9529	0,0471	70	0,9991	0,0009
17	0,9968	0,0032	44	0,9962	0,0038	71	0,9347	0,0653
18	0,9841	0,0159	45	0,9986	0,0014	72	0,9856	0,0144
19	0,9442	0,0558	46	0,9739	0,0261	73	0,9713	0,0287
20	0,9735	0,0265	47	0,9984	0,0016	74	0,9765	0,0235
21	0,9815	0,0185	48	0,9858	0,0142	75	0,9994	0,0006
22	0,9941	0,0059	49	0,9997	0,0003	76	1,0000	0,0000
23	0,9726	0,0274	50	0,9755	0,0245	77	0,9525	0,0475
24	0,9841	0,0159	51	0,9952	0,0048	78	0,9763	0,0237
25	0,9987	0,0013	52	0,9670	0,0330	79	0,9927	0,0073
26	0,9798	0,0202	53	0,9868	0,0132	80	0,9891	0,0109
27	0,9821	0,0179	54	0,9762	0,0238			

Conforme o Quadro 6, pode-se observar que as comunalidades obtidas são relativamente altas. Assim, o percentual da variância total atribuído aos 17 fatores comuns é grande e, conseqüentemente, o erro devido a aleatoriedade ou casualidade é pequeno.

A escolha das variáveis para definir cada fator pode ser obtida de acordo com a matriz de carregamentos fatoriais (ANEXO 05). Assim, as variáveis que apresentam altos carregamentos num mesmo fator podem ser agrupadas, procurando sempre estabelecer alguma relação técnica que justifique o agrupamento. A interpretação do agrupamento das variáveis nos fatores é subjetiva. A Rotação Varimax serve como ferramenta ao auxílio da interpretação, determinando os possíveis agrupamentos de acordo com a similaridade das variáveis em cada fator.

Os dados originais da matriz 20x80 foram reduzidos, sem grande perda de informação, e pelo Método dos Mínimos Quadrados Ponderados, a uma matriz de escores fatoriais \hat{f}_{ij} de ordem 20x17. Dispondo os fatores nas colunas e as observações (Faxinais) nas linhas, a matriz de escores fatoriais é apresentada no Quadro 7.

QUADRO 7 – MATRIZ DE ESCORES FATORIAIS

Faxinais	Fator1	Fator2	Fator3	Fator4	Fator5	Fator6	Fator7	Fator8	Fator9	Fator10	Fator11	Fator12	Fator13	Fator14	Fator15	Fator16	Fator17
1	-0,3791	-0,1873	-0,1411	0,0268	-0,2904	-0,1532	0,1135	0,053	-0,2306	0,0608	-0,2969	0,0112	0,1512	0,4311	-0,3441	0,423	0,1172
2	-0,6394	-0,6164	0,0212	0,2713	-0,993	-0,3145	0,2869	0,2756	-0,9006	-0,1342	-0,3913	-0,0751	0,2201	0,3895	-1,1771	0,8627	0,9933
3	-0,5047	-0,1522	0,0072	0,1838	-0,1516	-0,1262	0,0403	0,1724	-0,2593	0,0166	-0,0963	-0,1096	0,0984	0,1617	-0,1545	0,3324	0,2698
4	0,1082	-0,6896	-0,106	0,041	-1,2588	-0,2367	0,5274	-0,0205	-0,9146	-0,2213	-0,581	0,16	0,2931	0,4257	-1,5271	0,6813	1,0367
5	-0,6115	-0,012	-0,0442	0,1015	-0,1554	-0,1129	0,1567	0,1722	-0,2432	-0,0341	-0,0859	-0,029	0,0526	0,1951	-0,1928	0,3955	-0,1738
6	-0,2675	0,1132	0,0735	0,1418	0,1985	0,0062	-0,0568	0,2183	-0,0276	-0,0133	0,2157	-0,0663	-0,1522	-0,4079	0,2504	-0,1221	0,0693
7	-0,0315	0,5493	0,1954	-0,0249	0,7121	0,0699	-0,251	-0,0932	1,0307	0,266	0,1861	-0,1301	-0,0545	-0,2615	0,6688	-0,5657	-0,7563
8	-0,2606	-0,0234	-0,1222	-0,0045	-0,1347	-0,0769	0,1699	0,0445	-0,1399	-0,034	-0,1462	-0,0314	0,0702	0,2363	-0,2026	0,2485	-0,1521
9	0,0529	0,0859	0,0171	0,0428	0,5105	0,158	-0,5636	-0,0194	0,4731	0,1799	0,1189	-0,0179	-0,0354	-0,1481	0,6175	-0,2231	-0,2696
10	0,7934	-0,0942	0,074	-0,1785	-0,2237	0,1014	0,061	-0,1803	-0,0793	-0,1456	0,0477	0,1437	0,0147	-0,1266	-0,21	-0,2964	0,3569
11	0,3896	0,286	0,0771	-0,1473	0,5085	0,1425	-0,1109	-0,0834	0,2852	0,0778	0,3568	-0,0095	-0,1683	-0,3117	0,5751	-0,5432	-0,0482
12	-0,2206	-0,2345	-0,0915	0,0725	-0,3735	-0,1624	0,1492	0,0592	-0,292	0,0324	-0,2522	0,0074	0,1215	0,288	-0,4977	0,3413	0,333
13	-0,0477	0,3513	0,1529	0,0672	0,6918	0,1579	-0,2657	0,1151	0,3122	0,0495	0,4805	-0,0902	-0,2845	-0,6844	0,8096	-0,5181	-0,1951
14	-0,4176	0,2135	-0,1414	-0,0175	0,3626	-0,0313	-0,0921	0,0267	0,2489	0,1249	-0,0624	-0,0866	0,0241	0,2894	0,4439	0,1278	-0,5942
15	0,0351	0,106	0,0879	0,028	0,3404	0,0853	-0,2384	0,0433	0,1808	0,0785	0,235	-0,0589	-0,1129	-0,2395	0,4091	-0,179	0,0131
16	-0,7029	-0,6004	-0,1551	0,1793	-1,0983	-0,3808	0,5374	0,2379	-0,9188	-0,1447	-0,4909	-0,0292	0,2748	0,6049	-1,2798	0,9671	0,7037
17	2,3601	-0,01	0,2276	-0,3811	0,2874	0,5504	-0,4705	-0,6536	0,5553	-0,1062	0,4045	0,331	-0,1775	-0,6995	0,3857	-1,3153	0,3628
18	-0,1382	-0,0204	0,0655	0,0558	0,0637	-0,0083	0,0432	0,1311	-0,1017	-0,0224	0,0876	-0,0571	-0,0862	-0,161	0,0636	-0,0376	0,18
19	0,8176	0,5008	0,0298	-0,3248	0,753	0,2588	-0,1558	-0,3265	0,6542	0,0554	0,3481	0,0726	-0,2057	-0,2452	0,9675	-0,7984	-0,6293
20	-0,3357	0,4343	-0,2276	-0,1334	0,2509	0,0727	0,1192	-0,1721	0,3671	-0,086	-0,0779	0,0651	-0,0435	0,2638	0,3944	0,2191	-1,6173

Os escores médios finais de cada Faxinal foram calculados através de uma média dos escores fatoriais, ponderada pelos autovalores da matriz de correlação. Quanto maior é o autovalor, maior é a porção da variância total explicada por ele. Assim, no ranqueamento, os Faxinais mais bem classificados foram os que apresentaram escores fatoriais mais altos nos maiores autovalores.

Para atender às exigências do IAP, de acordo com a Portaria N.º 263/98, os escores brutos foram padronizados de forma que todos se situem no intervalo de 0 a 2. O Quadro 8 a seguir fornece o ranqueamento dos Faxinais analisados em função dos escores brutos e padronizados.

QUADRO 8 - RANQUEAMENTO DOS FAXINAIS ANALISADOS

Faxinal	Escore Bruto	Escore Padronizado
1º São Pedro	406,0385	1,8000
2º Ivai-Anta Gorda	255,5732	1,4586
3º Ponte Nova	163,4198	1,2495
4º Linha Brasília	150,4757	1,2201
5º Patos Velhos	116,3835	1,1427
6º Papanduva de Baixo	92,0459	1,0875
7º Queimadas	71,1442	1,0401
8º Cachoeira do Palmital	62,966	1,0215
9º Rio dos Couros	2,5779	0,8845
10º Rio do Meio	-9,2271	0,8577
11º Tijuco Preto	-11,4835	0,8526
12º Paraná-Anta Gorda	-19,0331	0,8355
13º Guanabara	-71,2905	0,7169
14º Salto	-101,915	0,6474
15º Taboãozinho	-105,585	0,6391
16º Dos Mellos	-113,344	0,6215
17º Marmeleiro de Baixo	-123,331	0,5988
18º Dos Krieger	-202,28	0,4197
19º Marmeleiro de Cima	-264,043	0,2795
20º Rio Bonito	-299,092	0,2000

4 RESULTADOS E CONCLUSÕES

Considerando-se a revisão detalhada da Análise Fatorial, a importância cultural, social e ecológica dos Faxinais, e a construção do programa automático que efetuou o ranqueamento das Áreas Especialmente Protegidas, pode-se concluir que :

- a) a amostra de dados brutos, obtida a partir dos questionários, foi suficiente para se efetuar o ranqueamento, considerando a precisão da estimativa e o grau de confiança da estimativa;
- b) os resultados dos escores finais obtidos com o programa, retratam com fidelidade a qualidade de cada Faxinal;
- c) as variáveis apresentaram-se moderadamente correlacionadas, fato que justificou a utilização da Análise Fatorial;
- d) a Rotação Varimax, utilizada para uma melhor interpretação dos dados, mostrou-se necessária em função do grande número de variáveis correlacionadas;
- e) a Análise Fatorial utilizada no ranqueamento cumpriu com os objetivos propostos;
- f) as variáveis utilizadas, atendem as exigências do Decreto Estadual N.º 3.446/97;
- g) o critério adotado em relação ao número de fatores (2.4.5) mostrou-se adequado para a análise;
- h) os 17 fatores considerados explicaram 97,7649% do total da variância;
- i) os valores altos das comunalidades mostraram que a maior parte da variabilidade apresentada pelas variáveis é devida aos fatores comuns;
- j) os Faxinais mais bem classificados, foram aqueles que apresentaram maiores coeficientes nos fatores associados a maiores autovalores;

- k) os Métodos Householder e QL mostraram-se confiáveis e eficientes na obtenção dos autovalores e autovetores da matriz de correlação;
- l) quando se objetiva uma redução de um número grande de variáveis, em um número bem menor com pouca perda de informação, estando as variáveis relativamente correlacionadas, a Análise Fatorial mostra-se muito confiável;
- m) programa atendeu as exigências da análise.

Tendo em vista este trabalho, sugere-se a implementação do programa no Sistema Computacional Geral do IAP e órgãos responsáveis pela questão.

ANEXOS

ANEXO 01 – QUESTIONÁRIO DO FAXINAL

QUESTIONÁRIO DO FAXINAL

1. Denominação do Faxinal:

2. Localização:

3. Distância em quilômetros do Faxinal até a sede do município:

4. Condições da principal estrada do Faxinal:

- Boas condições
 Condições regulares
 Más condições

5. O Faxinal conta com linha de transporte coletivo regular até a sede do município?

- Sim Não

6. N.º total de famílias _____

7. N.º total de habitantes _____

8. N.º de famílias de proprietários _____

N.º de arrendatários _____

N.º de famílias de proprietários/arrendatários _____

N.º de famílias em outra situação _____

9. Área

Total _____

Agricultura _____

Criadouro Comunitário _____

Reflorestamento _____

Florestas _____

Outros _____

10. Atividades realizadas no Criadouro Comunitário desde 1990:

- Reflorestamento com Erva-Mate
- Reflorestamento com Pinheiro do Paraná
- Melhoramento do Rebanho
- Retirada de Madeira

11. Qual o estado de conservação das cercas do Criadouro Comunitário ?

- Boas condições
- Condições regulares
- Más condições

12. Qual o comprimento da Cerca ?

13. Qual o perímetro do Criadouro Comunitário?

14. Quanto ao cadastro no IAP no que se refere ao ICMS Ecológico:

- É cadastrado
- Possui diretrizes conforme prevê o decreto estadual n,º 3,446/97
- Possui plano de ação negociada entre a Prefeitura e a comunidade

15. Possui Associação Comunitária:

- Sim
- Não

16. A associação possui sede própria:

- Sim
- Não

17. Qual o número de associados ?

18. A Associação oferece atividades próprias para as mulheres ?

- Sim
- Não

19. Possui outras organizações comunitárias ?

- Sim
- Não

20. Se positivo, quais ?

21. Através de uma escala de valores apresentada abaixo, assinale os equipamentos comunitários que possui a Associação e o seu estado de conservação

- (1) boas condições (2) condições regulares
(3) más condições (4) desativado

- () Viveiro () Oficina de Costura
() Armazém () Caminhão
() Descascador/Triturador () Outras máquinas agrícolas
() Cozinha () Outros

22. O Faxinal possui escola:

- () 1º Grau pública até 4ª série () Escola Particular
() 1º Grau pública até 8ª série () Não possui escola
() 2º Grau Pública

23. Se não possui escola, qual a distância, em quilômetros, da escola mais próxima ?

24. Quais os meios de transporte mais utilizados pelos estudantes?

25. O Faxinal possui Posto de Saúde Pública ?

- () Sim () Não

26. Se positivo, qual a regularidade do funcionamento?

- () 1 a 2 dias por semana
() 3 a 4 dias por semana
() todos os dias da semana

27. Qual a qualidade do atendimento do posto de saúde, se existente ?

- () De acordo com o desejado
() Inferior ao desejado

28. Possui telefone público comunitário?

- () Sim () Não

29. Possui Igrejas? Quais ?

30. Quanto ao potencial turístico, possui belezas cênicas:

- Cachoeiras
- Lagos
- Rios com corredeiras
- Outros _____

31. Qual o tamanho médio das propriedades rurais do Faxinal ?

32. Quais têm sido os principais problemas encontrados para a manutenção das cercas do Criadouro Comunitário?

- Problemas econômicos (falta de dinheiro)
- Problemas comunitários (falta de participação)
- Outros

33. Qual o tipo de abastecimento de água mais usado na comunidade ?

- Rede Pública
- Rede Comunitária
- Fonte Superficial (mina, olho d'água)
- Córrego ou rio
- Poço com sarilho
- Poço com bomba elétrica
- Poço com bomba manual

34. Que outras estruturas de produção existem dentro do criadouro comunitário ?
(assinalar conforme indicação abaixo, se é particular ou comunitária)

- tanques de peixe
 - plantações de fumo
 - estufa para secagem de fumo
 - barbaquás
 - cercas internas
 - outras atividades
- (1) particular (2) comunitária

ANEXO 02 – QUESTIONÁRIO UNIDADE FAMILIAR

QUESTIONÁRIO UNIDADE FAMILIAR

1. Nome do Faxinal:

2. A família do produtor mora na comunidade ?

() Sim () Não

3. Qual a área de terra que o agricultor possui ? Não esquecer a unidade de medida,

4. Com relação a terra que utiliza, o agricultor é:

() proprietário

() Posseiro

() Parceiro

() Arrendatário

() Outro _____

5. Qual o documento que tem da terra ?

() Título definitivo (matrícula)

() Título provisório (Escritura sem registro)

() Contrato de compra e venda

() Recibo

() Partilha

() Compromisso de compra e venda

() Documentos de herança em nome dos pais ou avós

() Outro _____

6. Como está sendo ocupada a terra, quanto à utilização?

(Não esquecer de colocar a unidade de medida):

A – Área explorada somente pelo proprietário _____

B – Área arrendada para terceiros _____

C - Área em parceria _____

D - Área utilizada para Criadouro Comunitário _____

E – Outra ocupação _____

7. O produtor se utiliza de outras áreas para produção (Arrenda ou faz Parceria) ?

() Sim () Não

8. Se utiliza, qual o tamanho da área (não esquecer a unidade de medida):

9. Renda anual total da unidade familiar (R\$):

14, Freqüência de consumo de alimentos pela família:

Tipo	1 = todos os dias 2 = uma vez por semana 3 = duas ou mais vezes por semana 4 = de vez em quando
	Leite
Derivados do leite	
Ovos	
Carne de frango	
Carne de gado	
Carne de porco	
Feijão	
Arroz	
Batata/mandioca (aipim)	
Farinha de trigo (pães e massas)	
Farinha de milho (pães e polenta)	
Farinha de mandioca	
Hortaliças de folhas	
Outras hortaliças	
Frutas	
Doces e bolos	
Açúcar	
Banha	
Óleo	
Outros produtos	

15, Como é a exploração dos recursos da floresta nas suas terras:

Espécie florestal	Produto ou uso	Quantidade	Plantada (P) Nativa (N)	

16, O produtor tem replantado mudas de árvores em suas florestas ?

Espécie florestal	N,º total de mudas		Área plantada (ha ou alq.)	
	Dentro do Criadouro	Fora do Criadouro	Dentro do Criadouro	Fora do Criadouro

20, Quais os produtos (insumos) abaixo utilizados no estabelecimento na safra ?

Tipo de insumo	Onde adquiriu (comércio, cooperativa, etc,)	Quantidade de área	Valor total gasto /unidade de medida
Adubação orgânica/verde			
Adubação química			
Calcáreo			
Fungicidas			
Herbicidas			
Inseticidas			
Mudas selecionadas			
Outros			

21. Foi utilizado crédito nas últimas três safras ?

() Sim () Não

Por que utilizou? _____

22. Quem presta assistência técnica?

() EMATER

() Associação

() Cooperativa

() Empresa de fumo

() Não recebe

() Outros _____

23. Há quanto tempo recebe assistência técnica?

24. Quais os tipos de culturas que são atendidas pela assistência técnica ?

25. O que acha da assistência técnica que recebe ?

() Ótima

() Satisfatória

() Regular

() Insuficiente

26, Como transporta os produtos até o local de venda?

27. Produção e comercialização por tipo de cultura agrícola na última safra – de julho de 1997 a junho de 1998 (não esquecer a unidade de medida):

Cultura	Área plantada	Área colhida	Quant, Colhida	Quant, Vendida	Valor total recebido

28. Para quem vendeu a última safra – de julho de 1997 a junho de 1998?

Produto	Comprador(es)	Município
_____	_____	_____
Produto	Comprador(es)	Município
_____	_____	_____
Produto	Comprador(es)	Município
_____	_____	_____
Produto	Comprador(es)	Município
_____	_____	_____
Produto	Comprador(es)	Município
_____	_____	_____

29. Plantel da última safra – de julho de 1997 a junho de 1998:

	N,º de cabeças em junho de 1998	Consumo próprio	N,º de cabeças compradas no período	N,º de cabeças vendidas no período	Valor total da venda (R\$)	Para quem vendeu	Gastos com rações (R\$)	Gastos com vacinas (R\$)
Bovino p/ corte								
Bovino p/ leite								
Bezerro p/ corte								
Bezerro p/ leite								
Animais p/ trabalho								
Suíno de raça								
Suíno comum								
Galinha								
Outras aves								
Outros								

30. Outros produtos da última safra – de julho de 1997 a junho de 1998:

Produto	Produção	Consumo próprio	Quantidade vendida	Valor da venda (R\$)	Quem comprou

31. De qual(ais) organizações sociais (associação, clubes, ONGs, etc.) o agricultor participa e há quanto tempo ?

_____	_____
_____	_____
_____	_____

32. Participa de atividades comunitárias? Quais?

- Viveiro
- Multirões
- Cozinha e costura
- Cerca
- Outros _____
- Não participa

33. Quanto tempo dedica às atividades comunitárias por mês?

34. Utilizando-se da escala abaixo, defina o(s) tipo(s) e condições das estradas e caminhos para escoamento da safra:

- (1) Boas condições (2) Condições regulares (3) Más condições
- Asfalto
- Saibro ou macadame
- Leito natural
- Picada
- Outros

35. Distância, em quilômetros, entre sua habitação e os seguintes serviços:

Escola primária _____ Sede do município _____
 Posto de Saúde _____ Asfalto _____

36. Qual o tipo de abastecimento de água usado na habitação?

- Poço de corda
- Poço de bomba elétrica
- Poço de bomba manual
- Fonte superficial (mina, olho d'água, etc,)
- Córrego ou rio
- Rede pública
- Outro

37. Qual a origem da iluminação existente na habitação ?

- Luz elétrica – gerador próprio
- Luz elétrica – rede pública
- Lâmpião a querosene ou diesel
- Lâmpião a gás
- Iluminação a gás
- Outro _____

38. Qual o tipo de banheiro existente na habitação?

- Banheiro dentro de casa
- Banheiro fora de casa
- Casinha
- A céu aberto
- Outro _____

39. Qual o destino do esgoto?

- () Fossa séptica
 () Fossa negra
 () Esgoto a céu aberto
 () Superficial
 () Outro _____

40. Normalmente a quem a família recorre para a prestação de assistência médica?

- () A própria família resolve o problema
 () Curandeiro ou benzedeira
 () Farmacêutico
 () Posto de saúde
 () Sindicato
 () Médico particular
 () Outro _____

41. De que material é construída sua casa?

- () Casa de alvenaria
 () Casa de madeira
 () Casa mista
 () Casa de madeira sem assoalho
 () Outros _____

42. Qual o tamanho da sua casa (metros quadrados) ?

43. Outras benfeitorias na terra:

Tipo	Material	Estado de conservação			Área (m ²)
		Bom	Regular	Ruim	

44. Qual religião você adota ?

45. Qual a principal origem dos antepassados?

- () Ucraniana () Italiana () Brasileira
 () Polonesa () Alemã () Outra _____

46. Sempre morou nesta comunidade ?

Sim

Não

De onde vem?

Do mesmo Município

Do mesmo Estado

De outro Estado _____

De outro País _____

47. De onde obtém mais informações que ajudam a resolver os problemas da lavoura e criação?

Cooperativa

EMATER

Parentes

Associação

Padre

Televisão

Amigos

Rádio

Outros _____

48. Coloque os seguintes itens em ordem de importância para você (de 1 até 7):

Educação

Infra-estrutura pública (estradas, luz, água)

Assistência técnica

Situação de emprego

Saúde

Proteção ao meio ambiente

Associações comunitárias

49. Que tipo de atividade o agricultor acha que poderia ser desenvolvida no Faxinal para melhorar a renda familiar?

Artesanato

Turismo

Plasticultura

Melhoria genética da criação

Outra _____

50. Estaria disposto a receber um turista em sua casa, se isto melhorasse sua renda ?

Sim

Não

51. Na sua opinião, como deve ser utilizada uma terra com florestas:

Cortada para dar lugar à agricultura

Explorada para produção florestal

Protegida

Explorada parcialmente

Outra finalidade

52. Na sua opinião a floresta serve para:

- Para produzir madeira/lenha
- Para produzir erva mate
- para proteger as águas
- Para proteger os animais e o meio ambiente
- Para exploração de outros produtos

53. Você considera a proteção do meio ambiente necessária para o futuro?

- Sim
- Não

Por quê ? _____

54. Você acredita que a proteção da natureza seja compatível com o desenvolvimento ?

- Sim
- Não

Por quê ? _____

55. O que pensa da cidade ?

- Nunca pensou
- É melhor para viver
- É pior para viver
- É bom para a compra e venda de produtos
- Outro

56. Qual a expectativa quanto ao futuro de seus filhos ?

- Estudar e morar na cidade
- Casar e morar na cidade
- Casar e trabalhar na lavoura morando com os pais
- Casar e trabalhar na lavoura distante dos pais
- Outros _____

57. Já pensou alguma vez em desistir da vida de lavrador ?

- Sim
- Não

58. Se pensou, o que pretendeu fazer com a propriedade ?

59. Já pensou em acabar com o Faxinal ?

- Sim
- Não

Por quê ? _____

60. Se o faxinal acabar, o que acha que vai acontecer ?

- Vai melhorar sua produção
- Vai melhorar a produção da maioria
- Muitos deixarão de ser agricultores
- Vai haver desmatamento
- Vai piorar a situação da maioria
- Outro _____

61. Você acha que a criação solta na floresta é condição para a existência do Faxinal ?

Sim Não

Por quê ? _____

62. Relacione, por ordem de prioridade, o que poderia ser desenvolvido para melhorar as condições de vida da sua comunidade ,

1. _____
2. _____
3. _____
4. _____
5. _____

63. Utilizando-se da escala abaixo, defina a atuação das organizações governamentais e não governamentais no faxinal:

(1) Boa (2) Regular (3) Ruim (4) Desconhece

- Prefeitura
- EMATER
- IAP
- Associação
- Sindicato
- Cooperativa
- ING
- Instituto de florestas Tropicais
- Outras

ANEXO 03 – MATRIZ DE DADOS ORIGINAIS

Matriz de Dados

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Marmeleiro de Baixo	84,21	5,26	57,89	2265,47	51,45	6,1	9,68	1,92	0,3	4,89	3,34	47,37	5,26	52,63	5,26
2 Marmeleiro de Cima	66,67	16,67	33,33	2896,67	67,03	0	7,26	6,45	0,91	6,35	2,52	33,33	33,33	16,67	0
3 Salto	85,71	0	57,14	3483,33	70,22	6,38	8,3	0	3,33	6,78	5,04	100	28,57	28,57	14,29
4 Dos Krieger	80	10	30	2453,4	75,73	10,6	4,74	0	0	4,11	13,16	70	10	60	0
5 Dos Mellos	100	0	20	2999	90	10	6,88	2,42	0,55	5,87	3,39	60	20	100	20
6 Rio dos Couros	66,67	11,11	33,33	5340	76,26	0	6,69	2,5	0,71	3,9	1,71	44,44	11,11	66,67	0
7 Ponte Nova	100	0	14,29	4805,71	76,99	7,88	11,58	8,23	1,61	8,23	15,73	100	28,57	71,43	14,29
8 Guanabara	100	0	12,5	2803	69,21	11,3	4,93	2,5	0,6	3,3	5,09	37,5	0	12,5	12,5
9 Queimados	73,91	26,09	52,17	4743,65	73,77	1,89	10,22	2,32	1,21	5,56	1,69	34,78	13,04	52,17	13,04
10 Papanduva de Baixo	73,68	21,05	31,58	5009,11	77,41	3,99	9,38	5,74	0,73	6,21	5,67	63,16	0	47,37	15,79
11 Marcondes (Linha Brasília)	100	0	33,33	5360,92	78,04	1,32	6,52	5,73	2,9	5,66	4,27	83,33	16,67	50	8,33
12 Taboãozinho	100	0	16,67	2924,92	42,37	13,03	6,89	5,01	0,85	4,49	7,23	75	25	75	16,67
13 Patos Velhos	75	25	41,67	6344,17	95,51	1,66	8,97	2,46	1,51	4,94	0	58,33	8,33	41,67	8,33
14 Linha Paraná-Anta Gorda	86,36	9,09	40,91	2466,82	64,1	11,45	6,25	6,82	0,53	4,66	3,89	36,36	9,09	45,45	13,64
15 Cachoeira do Palmital	76	4	56	5164,48	63,83	15,09	10,2	7,58	2,43	7,12	6,55	80	40	40	4
16 Rio Bonito	100	0	66,67	1600	21,67	25	2,15	7,26	0	3,69	1,21	66,67	33,33	33,33	33,33
17 São Pedro	91,67	8,33	33,33	7978,67	46,54	39,37	7,06	5,97	4,02	5,62	6,45	83,33	16,67	75	8,33
18 Rio do Meio	77,78	11,11	44,44	4637	32,16	8,67	5,24	6,45	1,13	4,96	8,33	88,89	0	33,33	11,11
19 Linha Ivaí-Anta Gorda	66,67	0	38,89	4922,89	51,02	30,56	8,8	6,98	3,59	8,05	13,55	61,11	16,67	72,22	16,67
20 Tijuco Preto	72	5,33	37,33	2055,18	34,14	13,29	4,3	4,88	1,38	4,21	5,52	48	6,67	54,67	8

Matriz de Dados

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	0	10,53	36,84	84,21	84,21	63,16	47,37	31,58	5,26	47,37	39,75	68,42	5,26	13,17	0	36,84	465,57	10,53
2	0	0	16,67	100	66,67	66,67	50	16,67	50	50	12	16,67	0	11	16,67	16,67	50	0
3	0	0	14,29	128,57	100	85,71	28,57	42,86	14,29	57,14	7,92	28,57	14,29	8	0	14,29	10	0
4	0	10	80	90	100	60	30	10	10	40	650	20	0	0	0	0	0	0
5	20	0	80	100	40	80	60	60	0	60	3,33	60	0	29,75	0	20	20	0
6	0	11,11	11,11	100	88,89	88,89	33,33	22,22	11,11	33,33	21	88,89	33,33	58	11,11	11,11	100	55,56
7	14,29	0	57,14	100	42,86	85,71	28,57	57,14	0	71,43	138,83	85,71	71,43	28,57	0	28,57	350	28,57
8	12,5	0	50	100	50	37,5	37,5	37,5	0	50	109,25	100	62,5	46,88	0	0	0	12,5
9	8,7	8,7	17,39	95,65	47,83	56,52	39,13	26,09	8,7	69,57	70,36	47,83	13,04	27,38	21,74	13,04	463,33	13,04
10	0	26,32	26,32	89,47	63,16	36,84	42,11	47,37	10,53	15,79	212,22	47,37	31,58	42,25	0	0	0	26,32
11	8,33	8,33	33,33	100	0	50	41,67	33,33	0	41,67	203,64	91,67	75	36	0	16,67	500	58,33
12	8,33	33,33	66,67	100	83,33	91,67	66,67	66,67	25	58,33	111,73	91,67	33,33	22,92	0	8,33	500	83,33
13	16,67	8,33	41,67	91,67	58,33	33,33	25	8,33	16,67	50	82	41,67	25	30	8,33	8,33	100	8,33
14	18,18	27,27	22,73	90,91	68,18	86,36	54,55	45,45	4,55	27,27	19,89	81,82	59,09	21,5	13,64	0	0	40,91
15	16	12	40	96	48	84	44	28	16	52	93,94	64	52	40,8	0	4	500	48
16	0	33,33	33,33	100	0	66,67	33,33	0	0	100	43,33	100	0	23,33	0	0	0	0
17	16,67	0	58,33	100	50	58,33	41,67	58,33	16,67	58,33	563	91,67	8,33	27,55	0	0	0	66,67
18	22,22	0	11,11	100	44,44	33,33	22,22	44,44	0	55,56	172,38	88,89	33,33	62,44	0	11,11	80	22,22
19	22,22	38,89	33,33	100	66,67	72,22	55,56	66,67	5,56	0	237,94	100	38,89	39,56	0	5,56	50	55,56
20	6,67	10,67	49,33	86,67	58,67	72	44	50,67	4	54,67	123,98	76	25,33	19,97	0	4	200	30,67

Matriz de Dados

	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
1	350	21,05	133,25	84,21	31,58	0	5,26	63,16	465	742,08	3000	414,4	3,33	2,38	9,18	23,17	78,95	94,74
2	0	0	0	50	0	0	0	33,33	600	1810	0	0	4	2,75	10,17	15,17	83,33	83,33
3	0	0	0	42,86	14,29	0	28,57	57,14	394	1974	3000	0	5,5	3,33	17	30,71	85,71	85,71
4	0	0	0	20	30	0	10	70	710	80	0	1300	8	1,75	15,14	34,38	0	70
5	0	80	1250	60	20	0	0	100	60	1006,67	3100	0	4	4,25	4,758	33	100	100
6	350	66,67	900	66,67	55,56	22,22	0	88,89	120	135,83	4080,63	88	3,89	4,88	4,25	33,33	100	100
7	120	57,14	1025	100	85,71	14,29	0	85,71	655,33	2213,33	6500	476,67	8	3,14	8	57,14	85,71	57,14
8	700	37,5	766,67	87,5	50	25	0	62,5	96	322,5	3500	536,67	3,14	3	7,67	30,88	12,5	62,5
9	233,33	34,78	1300	65,22	21,74	39,13	13,94	39,13	2247,31	1656,38	5650	350	3	2,1	4,71	37,08	73,91	78,26
10	650	57,89	983,64	89,47	42,11	5,26	26,32	52,17	836,5	1521,33	3750	1848	4,76	3,47	13,64	32,94	57,89	89,47
11	821,43	50	416,67	75	66,67	16,67	8,33	75	0	1859,38	5686,42	1079,5	8,92	4,4	7,11	45,42	58,33	83,33
12	1055	33,33	250	58,33	25	33,33	0	25	515	751	2650	547	6,5	4,17	15,1	41	91,67	75
13	500	16,67	1000	58,33	33,33	0	16,67	58,33	313,33	115,56	5731,67	167,5	5,5	2,11	3	26	66,67	66,67
14	558,89	13,64	433,33	63,64	50	9,09	0	36,36	313,33	1214,1	5250	202,11	2,85	3,05	19,06	33,68	27,27	86,36
15	1226,67	28	985,71	80	60	28	16	44	834,11	2201,75	4814,29	444	5,77	3,86	41,93	73,96	76	76
16	0	33,33	15	33,33	66,67	100	0	0	0	400	0	200	10,5	2	0	35	100	33,33
17	4462,5	41,67	1840	50	41,67	16,67	8,33	50	2700	2440	6183,33	3428,55	7,83	4,75	10	47,27	83,33	41,67
18	225	33,33	406,67	77,78	33,33	0	0	44,44	203,25	701,67	3833,33	424,29	9,5	2,25	4,2	30	33,33	11,11
19	1466	11,11	915	77,78	44,44	16,67	0	22,22	163,33	1816,94	7250	1757,69	6,44	3,44	15,94	39,38	22,22	61,11
20	522,61	9,33	4698,57	28	10,67	2,67	1,33	26,67	348,89	455,39	5233,33	487	4,88	2,42	16,28	29,03	53,33	68

Matriz de Dados

	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
1	1,25	0,8	0,8	15	5,26	89,47	78,95	21,05	10,53	0	89,47	49,61	100	21,05	73,68	57,89	73,68	94,74
2	1,2	6	6	18	16,67	0	66,67	50	50	0	100	57,33	66,67	50	83,33	33,33	83,33	100
3	0,93	1,5	10	15	71,43	0	85,71	28,57	28,57	0	85,71	82,44	128,57	57,14	114,29	85,71	28,57	100
4	2	0,1	9	11	20	0	80	20	10	0	80	51	70	0	90	80	80	100
5	0,75	2,5	3	9	20	0	60	40	20	0	80	66,3	80	20	100	20	60	100
6	1,78	5	5,5	15	77,78	0	100	11,11	22,22	0	88,89	68,56	100	66,67	77,78	22,22	22,22	100
7	1	0,1	1	5	100	0	100	42,86	57,14	14,29	71,43	81,14	57,14	0	0	42,86	85,71	71,43
8	0,58	0,05	4	3	37,5	0	87,5	0	0	25	0	46	62,5	25	50	62,5	100	100
9	1,87	0,2	10	7	8,7	78,26	62,61	34,78	17,39	17,39	65,22	70,67	82,61	47,83	69,57	21,74	43,48	86,96
10	0,9	0,4	1	10	47,37	0	78,95	36,84	31,58	15,79	63,16	85,55	84,21	68,42	94,74	42,11	68,42	73,68
11	0,8	0,2	0,2	0,1	25	16,67	100	0	0	41,67	58,33	64,21	75	16,67	100	33,33	75	100
12	0,98	0,3	4	8	66,67	0	100	25	16,67	33,33	66,67	76,33	75	50	83,33	41,67	66,67	100
13	2,88	0,1	0,2	4	33,33	16,67	50	50	25	16,67	83,33	64	75	25	91,67	58,33	33,33	100
14	0,89	2	2,5	2,5	22,73	0	54,55	18,18	9,09	36,36	50	57	81,82	68,18	81,82	77,27	86,36	100
15	1,83	0,5	0,6	6	28	16	80	36	24	40	32	71,9	76	32	92	40	80	92
16	1,5	0,5	26	8	33,33	0	33,33	33,33	0	0	66,67	47	66,67	0	66,67	66,67	0	66,67
17	0,65	0,05	4	2	16,67	83,33	100	66,67	41,67	58,33	0	89	91,67	75	66,67	16,67	41,67	75
18	1	0,1	21	1	55,56	0	66,67	66,67	55,56	22,22	77,78	58,81	77,78	33,33	88,89	33,33	77,78	100
19	0,86	0,3	4	4	50	0	100	44,44	33,33	44,44	38,89	74,78	77,78	33,33	72,22	44,44	72,22	77,78
20	1,6	0,1	0,2	1	17,33	24	58,67	10,67	6,67	21,33	62,67	49,64	77,33	45,33	60	42,67	78,67	90,67

Matriz de Datos

	70	71	72	73	74	75	76	77	78	79	80
1	94,74	94,74	73,68	1	1	1	1	2	1	1,14	1
2	100	33,33	83,33	0	1	1	1	1	0	4,96	0
3	128,57	85,71	85,71	0	1	1	1	0	0	2,69	1
4	100	100	70	1	1	1	0	1	0	0,56	0
5	100	60	100	0	1	1	0	0	0	0,4	1
6	100	100	100	0	1	1	1	0	0	0,68	1
7	85,71	57,14	100	1	1	1	0	1	0	0	0
8	100	75	87,5	1	1	1	0	1	0	2,56	0
9	82,61	69,57	100	0	0	0	1	2	0	0	1
10	68,42	57,89	84,21	0	1	0	1	2	1	0,36	0
11	100	100	83,33	1	1	1	1	2	1	1,39	1
12	91,67	83,33	100	0	1	0	1	0	0	1,07	1
13	100	25	91,67	0	1	0	1	1	1	0	1
14	95,45	95,45	100	0	1	0	0	0	0	0,96	0
15	76	64	100	0	1	0	1	1	0	0,38	1
16	66,67	100	100	0	0	0	0	1	0	0	0
17	75	50	91,67	0	0	0	1	1	0	0	1
18	66,67	55,56	100	0	0	0	1	1	0	0	0
19	77,78	72,22	100	1	1	0	0	1	0	0,37	0
20	81,33	72	96	0	1	0	1	2	1	0,95	1

ANEXO 04 – MATRIZ DE CORRELAÇÃO

CONTINUAÇÃO

Matriz de Correlação

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18			
1	1,0000																				
2	-0,6118	1,0000																			
3	-0,3108	0,1661	1,0000																		
4	-0,1835	0,3291	-0,0818	1,0000																	
5	-0,0569	0,3306	-0,3482	0,2994	1,0000																
6	-0,1984	-0,4073	0,0780	0,1582	-0,5669	1,0000															
7	-0,2191	0,2628	-0,0063	0,4728	0,4908	-0,2421	1,0000														
8	0,0605	-0,1665	-0,0200	0,1791	-0,4087	0,3668	0,0500	1,0000													
9	-0,0670	-0,1977	0,0959	0,6577	-0,0092	0,4304	0,3364	0,2047	1,0000												
10	-0,1499	-0,0899	0,0001	0,3944	0,2449	0,0911	0,7367	0,3774	0,6009	1,0000											
11	0,0715	-0,3790	-0,3845	0,0676	-0,0968	0,3329	0,1895	0,2516	0,4747	0,4747	1,0000										
12	0,3735	-0,4134	0,0071	0,3400	-0,0985	0,2384	0,1532	0,1830	0,4984	0,4404	0,5101	1,0000									
13	0,1833	-0,3513	0,2036	-0,0421	-0,0708	0,1786	0,1583	0,3320	0,2623	0,4564	0,0545	0,3419	1,0000								
14	0,1731	-0,2022	-0,3545	0,2214	0,1267	0,2806	0,2031	-0,0086	0,1226	0,2147	0,3144	0,1872	0,0267	1,0000							
15	0,5216	-0,3295	0,1258	-0,2440	-0,3247	0,3655	-0,2191	0,2878	-0,0749	0,0162	-0,0456	0,1773	0,1678	0,1031	1,0000						
16	0,1250	-0,1202	-0,2273	0,3828	0,0146	0,3706	0,1519	0,3780	0,3562	0,2954	0,2707	0,1611	-0,0885	0,2867	0,1677	1,0000					
17	-0,0824	-0,0978	0,1878	-0,2142	-0,3621	0,3217	-0,1271	0,3335	-0,1158	-0,0520	0,0697	-0,1511	-0,0491	0,1522	0,4556	-0,0458	1,0000				
18	0,5077	-0,3634	-0,5372	-0,1271	0,1094	0,3286	-0,1166	-0,1427	-0,1111	-0,0827	0,3847	0,1873	0,0657	0,5620	0,0780	0,1838	-0,0523	1,0000			
19	0,2186	-0,3638	0,0692	0,1079	0,0466	0,0477	0,0152	-0,1492	0,4779	0,3050	0,0350	0,4944	0,4546	-0,1579	0,2309	-0,0336	-0,2420	-0,2384	1,0000		
20	-0,4405	0,1632	-0,0784	-0,1345	0,1613	-0,1730	0,2110	-0,5392	-0,0846	-0,0241	0,1784	-0,1352	-0,1802	0,0656	-0,4856	-0,2923	0,0007	-0,0008	0,2099	1,0000	
21	0,0969	-0,4803	-0,0087	-0,2633	-0,0790	0,1229	0,1229	0,1005	0,0681	0,0861	-0,1839	0,0975	0,6099	0,1757	0,1453	-0,0816	0,2099	0,1259	0,2099	0,2099	1,0000
22	0,1188	-0,2682	-0,2518	-0,2591	-0,1005	0,2148	0,0498	0,1991	-0,0194	0,0990	0,0406	-0,3355	0,1757	0,4131	0,1418	0,1453	0,4445	0,2812	0,2812	0,2812	1,0000
23	0,3027	-0,2134	-0,3681	0,1147	0,1652	-0,0627	0,4291	0,3406	0,0444	0,4801	0,5622	0,4056	0,2419	0,2261	0,0960	0,0960	-0,2169	0,0007	0,0007	0,0007	1,0000
24	-0,3905	0,3569	-0,0273	0,0798	0,0845	-0,1475	0,1708	0,0222	0,0751	-0,2025	-0,2504	0,2180	0,4273	-0,1111	-0,1111	-0,1111	-0,2270	-0,4283	-0,2985	-0,1423	1,0000
25	0,5080	-0,1432	0,2095	-0,2128	-0,2809	0,0493	-0,2313	0,0136	-0,2025	0,2241	-0,2604	0,2180	0,4273	-0,1111	-0,1111	-0,1111	-0,2270	-0,4283	-0,2985	-0,1423	1,0000
26	0,0105	0,0351	-0,2227	0,3291	-0,0641	0,4693	-0,1473	-0,0548	0,2523	-0,0197	0,5446	0,3152	-0,2036	0,2576	-0,2585	0,0507	-0,0398	0,0384	0,0384	0,0384	1,0000
27	0,4113	-0,5268	-0,1564	0,1045	-0,4911	0,4836	-0,2547	0,4863	0,1411	-0,1641	0,1395	0,1044	-0,1148	0,2159	0,3997	0,4111	0,3239	0,0384	0,0384	0,0384	1,0000
28	0,2279	-0,2726	-0,3539	0,2396	0,1666	-0,1259	0,2148	0,4279	0,1229	0,2032	0,3002	0,1884	-0,0879	-0,0962	-0,0292	0,3849	0,0644	-0,1053	-0,4890	-0,4890	1,0000
29	-0,1126	0,0501	-0,1601	0,4617	-0,0167	-0,0202	0,0298	0,3178	0,0922	-0,0147	-0,10306	0,0354	-0,3128	-0,0066	0,0504	0,4353	0,0047	-0,3252	-0,3252	-0,3252	1,0000
30	-0,4470	0,6751	0,1431	0,0486	0,2916	-0,4258	0,1826	-0,0814	-0,2058	-0,0751	-0,4730	-0,6410	0,0025	-0,2009	-0,2567	-0,0983	-0,0603	-0,4890	-0,4890	-0,4890	1,0000
31	0,1269	-0,1277	0,0100	-0,0693	0,2151	-0,4341	0,4996	-0,1363	-0,0453	0,3398	0,0097	0,1092	0,1301	0,1992	-0,1518	-0,1009	-0,3829	-0,0462	-0,0462	-0,0462	1,0000
32	0,1357	-0,0886	0,0829	0,0769	-0,0115	-0,2572	0,4588	0,1264	0,0889	0,1839	0,0216	0,1297	0,2386	0,1552	-0,1691	-0,0207	0,0284	0,0405	0,0405	0,0405	1,0000
33	0,0550	-0,2315	-0,2647	0,4511	-0,2198	0,3461	0,1171	0,4337	0,4464	0,1116	0,2273	0,2003	0,0483	0,4046	-0,1016	0,2937	0,3798	0,0504	0,0504	0,0504	1,0000
34	0,0781	-0,0666	-0,1045	0,6321	-0,2261	0,1788	0,0744	0,2775	0,6377	0,1224	0,1297	0,1979	-0,0022	0,2788	-0,0876	0,3513	0,0303	0,2016	0,2016	0,2016	1,0000
35	0,3877	-0,1029	-0,3967	0,3219	0,2636	-0,0827	0,1291	0,1380	-0,1076	0,0206	-0,0888	0,1229	-0,0840	0,4889	0,2535	0,1740	-0,1186	0,1672	0,1672	0,1672	1,0000
36	-0,2089	0,0617	-0,1367	0,1144	-0,1572	0,1964	-0,0604	0,1012	0,1870	-0,0447	0,0041	-0,1496	-0,2299	0,2676	-0,0784	0,1971	-0,1113	0,1921	0,1921	0,1921	1,0000
37	0,0905	-0,0325	-0,2402	0,3417	0,2491	-0,1953	0,6110	0,3323	0,0974	0,4119	0,1870	0,0712	-0,1744	0,0042	0,0150	0,3489	-0,0397	-0,0397	-0,0397	-0,0397	1,0000
38	0,3906	-0,3265	-0,0693	0,2764	0,0265	0,1854	0,0988	0,4766	0,0857	0,1307	0,2785	0,3366	0,1242	0,0812	0,2360	0,1923	0,1945	0,0231	0,0231	0,0231	1,0000
39	0,3316	-0,2066	0,3175	-0,1588	-0,4455	0,3493	-0,3300	0,3977	-0,1508	-0,2540	-0,2110	-0,0221	0,3927	-0,1069	0,6133	-0,1583	0,4657	-0,0628	-0,0628	-0,0628	1,0000
40	-0,2147	0,3786	0,3519	0,3307	0,3841	-0,2159	0,4196	-0,3478	0,3131	0,2717	-0,1997	0,2940	0,0094	-0,2173	-0,0881	-0,2835	-0,1007	-0,2013	-0,2013	-0,2013	1,0000
41	0,1918	-0,0490	-0,4853	0,2562	0,7293	-0,4304	0,3026	-0,4244	-0,0408	0,1032	0,0915	0,1604	-0,1951	0,3563	-0,3697	0,0368	-0,6041	0,2667	0,2667	0,2667	1,0000
42	-0,1113	0,4251	0,0732	0,5030	-0,0095	0,3081	0,3478	0,0042	0,3383	0,1764	0,2738	0,0318	0,0410	0,1784	-0,1711	0,0442	-0,2039	0,0633	0,0633	0,0633	1,0000
43	0,0411	-0,1098	0,0549	0,4538	0,0835	0,2375	0,5768	0,4490	0,7070	0,8197	0,2035	0,1398	-0,2060	0,0912	0,0123	0,2186	-0,1307	-0,1498	-0,1498	-0,1498	1,0000
44	-0,0887	0,0386	-0,1183	0,6400	0,1636	0,1792	0,5042	0,2802	0,6111	0,4270	0,2035	0,1398	-0,2060	0,3273	-0,0262	0,6295	0,0021	-0,0891	-0,0891	-0,0891	1,0000
45	0,0073	0,0190	-0,1750	0,5527	-0,1477	0,6424	0,0516	0,2044	0,5119	0,1787	0,3833	0,2584	-0,1990	0,3057	-0,0564	0,1597	0,1242	0,2587	0,2587	0,2587	1,0000
46	0,3280	-0,3166	0,1059	0,1767	-0,4141	0,3915	-0,3199	0,3772	0,2366	0,0793	0,3961	0,7286	0,2621	0,0607	0,2921	0,0861	0,1010	0,1611	0,1611	0,1611	1,0000

CONTINUAÇÃO

Matriz de Correlação

	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
1																			
2	1,0000																		
3	0,0639	1,0000																	
4	0,2785	0,2955	1,0000																
5	-0,1854	0,0746	0,4408	1,0000															
6	0,0635	-0,1302	0,2483	-0,2405	1,0000														
7	0,1202	0,3821	0,1419	0,2378	-0,1986	1,0000													
8	0,2070	-0,4290	0,0803	-0,2742	0,2397	-0,0626	1,0000												
9	-0,1861	0,1088	-0,2744	-0,1740	-0,0085	-0,0374	-0,1915	1,0000											
10	-0,0584	-0,4544	0,0781	0,1619	0,1446	-0,5371	0,0459	-0,0434	1,0000										
11	0,0018	-0,2692	0,0168	-0,0293	0,4085	-0,3449	-0,2761	-0,1174	0,5036	1,0000									
12	-0,0141	-0,3205	-0,2937	-0,1724	-0,0102	-0,3333	-0,2178	-0,1543	0,5600	0,4845	1,0000								
13	-0,1042	0,1159	0,0417	0,0369	-0,1292	0,4170	-0,0159	-0,3361	-0,3410	-0,1298	-0,0680	1,0000							
14	0,0552	0,0425	0,1891	0,0251	0,4296	-0,0184	0,1731	-0,3826	-0,1123	-0,0342	-0,1469	0,0339	1,0000						
15	-0,1886	-0,1541	0,2016	0,2073	0,2073	-0,0028	0,1849	-0,1896	0,1375	0,3203	-0,0023	0,0278	0,4853	1,0000					
16	-0,0520	-0,0179	0,2919	0,4032	0,0130	0,0384	-0,3267	0,2028	0,5932	0,4916	0,3651	-0,1428	0,1855	-0,1927	1,0000				
17	-0,0206	-0,0791	-0,0617	0,1991	-0,1220	0,1108	-0,1460	0,5371	0,3660	0,0545	0,1373	-0,2079	-0,3319	-0,3319	-0,3319	1,0000			
18	-0,0326	-0,4156	0,0005	0,0686	0,2687	-0,3912	0,1193	-0,1233	0,4110	0,2347	0,5799	-0,1401	0,1855	0,0254	0,3756	1,0000			
19	-0,0319	-0,1031	0,0125	0,0710	0,0409	-0,1883	0,0065	0,0677	0,1649	0,0432	0,0957	-0,0702	-0,1787	0,0364	0,1819	0,2338	1,0000		
20	-0,0869	-0,1754	-0,1617	0,0425	0,4051	-0,2461	-0,2879	-0,2842	0,3539	0,6352	0,5996	-0,0676	0,3745	0,3408	0,2479	0,0847	0,0847	1,0000	
21	-0,0573	-0,4801	0,0491	-0,2467	0,5097	-0,5156	0,0516	0,0781	0,5999	0,6512	0,4185	-0,2714	-0,0374	0,1727	0,3271	0,2795	0,2795	0,2795	1,0000
22	0,0793	-0,5202	0,1245	-0,0039	-0,0265	-0,1745	0,5350	-0,1432	-0,4509	-0,0331	0,0865	0,0152	-0,2944	0,1098	0,1727	0,1098	0,1098	0,1098	0,1098
23	0,2496	0,2308	0,0113	-0,2018	-0,1748	-0,1335	-0,1229	0,1084	-0,5731	-0,1153	-0,1688	-0,0336	-0,1162	0,0128	0,0128	0,0128	0,0128	0,0128	0,0128
24	0,0662	0,1312	0,0876	-0,0159	0,3212	-0,2368	-0,1221	0,0601	-0,1534	0,1727	0,1522	-0,0998	-0,1230	0,0337	0,0337	-0,1101	-0,1346	-0,1346	-0,1346
25	-0,0676	0,0602	-0,0876	-0,0159	0,0272	0,2539	0,1531	0,4525	-0,1719	-0,2581	-0,1587	0,2839	-0,1230	0,1360	0,1360	0,1771	0,1771	0,1771	0,1771
26	0,3607	-0,1814	0,2217	0,1802	0,3156	0,2084	-0,0672	0,0752	-0,0417	0,2245	-0,0667	0,0148	0,1727	0,2331	0,2331	0,2795	0,2795	0,2795	0,2795
27	-0,0676	-0,1686	-0,0203	0,0171	0,2865	-0,3439	-0,3202	0,0231	0,4109	0,5789	0,4173	0,0159	0,0519	0,2675	0,2675	0,5211	0,5211	0,5211	0,5211
28	-0,1378	-0,0427	-0,2731	-0,0711	-0,0532	-0,0228	-0,3262	0,8034	0,2064	-0,0065	0,0617	-0,3353	-0,3663	-0,1606	-0,1606	0,4532	0,4532	0,4532	0,4532
29	0,1871	-0,4560	-0,1440	-0,3766	0,2197	-0,2152	0,3331	0,4535	0,2900	0,0437	0,0679	-0,5219	-0,1401	-0,0045	-0,0045	0,1454	0,1454	0,1454	0,1454
30																			
31																			
32																			
33																			
34																			
35																			
36																			

Matriz de Correlação

CONTINUAÇÃO

	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		
18																		
19																		
20																		
21																		
22																		
23																		
24																		
25																		
26																		
27																		
28																		
29																		
30																		
31																		
32																		
33																		
34																		
35																		
36																		
37	1,0000																	
38	0,4940	1,0000																
39	-0,1253	0,4348	1,0000															
40	-0,0218	-0,1491	-0,1982	1,0000														
41	0,3154	0,1738	-0,5011	0,0851	1,0000													
42	-0,0645	-0,1372	0,0270	0,3078	-0,0910	1,0000												
43	0,3253	0,1487	-0,0824	0,2906	-0,0179	0,4824	1,0000											
44	0,4924	0,3337	-0,1985	0,0635	0,1390	0,2502	0,4106	1,0000										
45	0,0294	0,1561	-0,0600	0,1644	-0,0698	0,5716	0,3760	0,3180	1,0000									
46	-0,1830	0,4098	0,3532	-0,0789	-0,1871	-0,0628	0,0766	-0,0938	0,3073	1,0000								

Matriz de Correlação

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
47	0,2069	-0,3440	-0,3741	0,4664	0,1838	0,1748	0,1771	0,1680	0,4544	0,2313	0,0020	0,2329	0,2111	0,4655	-0,0798	0,1680	-0,0122	0,0929
48	-0,2221	-0,1879	0,1443	-0,0041	-0,0399	0,1754	0,2765	0,2118	0,2943	0,3549	0,2896	0,1616	0,3448	-0,0624	-0,2792	0,1081	-0,1762	0,0711
49	0,2562	-0,3337	-0,0373	0,4102	0,0138	0,3365	0,3522	0,4443	0,4253	0,4397	0,4238	0,5345	0,4607	0,2888	0,0523	0,3539	0,0730	0,2508
50	0,1783	-0,0482	0,1785	0,1300	0,0043	-0,0868	0,2392	0,0744	0,0515	0,1616	-0,4260	0,1583	0,5633	0,2592	0,2003	-0,2532	-0,1043	-0,0719
51	-0,1452	0,0855	-0,1203	-0,1563	-0,1633	-0,5157	0,3579	-0,4172	-0,1526	0,0781	-0,3064	-0,3888	0,0324	0,2077	-0,3067	-0,3714	0,0134	0,0043
52	-0,4384	0,5595	0,3590	0,0831	0,2217	-0,3701	-0,0859	-0,2854	-0,2369	-0,2122	-0,3731	-0,1969	0,0220	-0,1055	-0,3137	-0,1838	-0,0103	-0,0454
53	-0,3439	0,1448	-0,0503	-0,1436	0,2294	-0,3710	-0,0387	0,0061	-0,2158	-0,0064	-0,3731	-0,4186	0,2532	-0,0873	-0,3394	-0,3152	-0,1680	-0,3545
54	0,0782	-0,0260	0,4442	-0,2582	-0,5310	0,1810	-0,5332	0,0396	-0,2334	-0,2903	-0,0904	0,1514	0,1217	-0,3058	0,4473	-0,1363	0,0949	-0,2895
55	-0,2747	0,1382	0,1757	-0,2987	0,2077	-0,3798	0,1726	-0,4252	-0,3139	0,0391	-0,2187	-0,1959	0,2771	-0,0943	-0,2374	-0,7070	-0,0719	-0,1854
56	0,1180	-0,2786	-0,3072	0,1783	0,0542	-0,1017	0,2072	0,1561	0,1313	0,2918	0,4234	0,5175	0,1623	0,1061	0,1690	0,0264	0,0565	-0,1247
57	-0,0512	0,2691	0,3343	0,2788	-0,1244	0,1788	0,3041	-0,1737	0,2079	-0,0365	-0,2388	-0,1584	-0,1648	0,1492	-0,2017	-0,0401	-0,2079	-0,0043
58	0,0912	-0,3427	-0,4705	0,4323	0,1483	0,1091	0,3630	0,0055	0,4898	0,3488	0,5257	0,3742	0,0242	0,3132	-0,3446	0,0239	-0,0844	0,1326
59	-0,2043	0,3173	-0,1403	0,4658	-0,1367	0,3341	0,2700	0,3333	0,2900	0,4511	0,1548	0,3185	0,1897	0,1219	0,1100	0,4118	-0,1397	-0,0622
60	-0,2893	0,2200	-0,1477	0,4743	0,0384	0,0339	0,4590	0,3536	0,3229	0,6297	0,3903	0,3822	0,1739	0,0764	-0,1700	0,2855	-0,3190	-0,1974
61	0,1051	-0,1405	-0,1510	0,5281	-0,2078	0,5470	0,1070	0,5107	0,6319	0,2133	0,2485	0,1928	-0,0353	0,1499	0,0038	0,6156	0,2196	0,0840
62	-0,0265	0,0746	-0,1559	0,6898	0,2444	-0,6570	0,0818	-0,2962	-0,4186	0,0126	-0,1896	-0,0351	0,0926	0,0197	-0,1434	-0,4531	-0,0902	-0,2259
63	-0,2163	0,0200	0,4350	0,1319	0,0439	-0,0676	0,1832	-0,4960	0,3281	0,0572	-0,2791	0,1387	-0,0665	0,0434	-0,1265	0,0619	0,0114	-0,3804
64	-0,3969	0,3491	0,0158	0,3170	-0,0328	0,0398	0,1267	0,0207	0,2727	0,0027	-0,2993	-0,2305	-0,1718	-0,0112	-0,2038	-0,0332	0,0764	-0,4308
65	-0,2286	0,1791	0,3389	0,0116	0,1726	-0,2060	-0,1375	-0,3711	0,0857	-0,1002	-0,4113	0,0227	-0,0201	-0,1080	-0,1215	-0,1349	0,0741	-0,2014
67	0,1296	-0,1907	0,2593	-0,5239	-0,0366	-0,0319	-0,2269	-0,2887	-0,2141	-0,2026	0,1033	0,0561	-0,0301	-0,0441	0,1442	-0,2725	0,2166	0,0095
68	0,0165	-0,1526	-0,4968	-0,1902	0,0856	-0,1274	0,1434	0,2178	-0,0805	-0,1752	-0,2874	-0,1095	-0,2385	-0,0860	-0,3552	0,2907	-0,1455	-0,0350
69	-0,0884	0,0502	-0,1516	-0,2139	0,3252	-0,5495	-0,1484	-0,5236	-0,1977	-0,3441	-0,1842	-0,0546	0,0841	-0,0998	-0,3101	-0,2736	-0,3455	-0,0350
70	0,1135	-0,1417	-0,1535	-0,1963	0,5465	-0,4630	0,0672	-0,6702	0,0315	-0,0524	-0,1842	-0,0546	0,2513	-0,0898	0,3101	-0,2736	-0,3733	0,0328
71	0,2658	-0,4692	0,1967	-0,4573	-0,2299	0,0209	-0,3760	-0,2269	-0,2207	-0,3901	0,0312	-0,0177	-0,0462	0,0758	0,1229	-0,3986	0,3670	-0,0154
72	0,0460	-0,1126	-0,0128	0,1688	-0,2337	0,2513	0,0250	0,4991	0,1216	0,1468	-0,0295	0,0420	0,2513	0,2717	0,4547	0,5641	0,2493	-0,1329
73	0,2470	-0,3960	-0,2850	-0,0980	0,1426	0,0269	0,1157	-0,0879	0,0477	0,1341	0,5512	0,0938	-0,1821	0,0529	-0,1426	0,0014	-0,0340	0,2965
74	-0,0819	-0,2209	-0,3924	-0,2251	0,5012	-0,3901	0,2467	-0,1885	-0,0747	0,1749	0,1572	-0,0182	0,0152	0,0709	-0,3438	-0,3788	0,0573	0,2140
75	0,2351	-0,3074	-0,3157	-0,2237	0,4641	-0,4426	0,0377	-0,4559	-0,1558	0,0164	0,0682	0,0182	0,0728	-0,0213	-0,3399	-0,3438	-0,5426	-0,4200
76	-0,3928	0,4279	0,2825	0,3919	-0,0468	-0,3191	0,2637	-0,0961	0,2698	0,0078	-0,4067	0,0678	-0,0435	-0,1743	-0,4398	-0,2625	-0,2172	-0,4200
77	-0,1999	0,3142	0,2121	0,0972	-0,1031	-0,1002	0,1566	0,1090	0,0331	0,0434	-0,0132	-0,1335	-0,3151	-0,1984	-0,1307	-0,1978	-0,0405	-0,1069
78	-0,1326	0,2513	0,1090	0,0694	0,1328	-0,3216	0,1175	-0,0962	-0,0248	-0,1060	-0,2739	-0,0966	-0,4309	-0,0596	-0,1533	-0,2270	0,0414	-0,0410
79	-0,0920	-0,0818	-0,1228	-0,3587	0,0970	-0,3310	-0,1066	-0,1496	-0,0283	-0,0052	-0,2080	-0,2893	0,2064	-0,5458	-0,3177	-0,3961	-0,2776	-0,2413
80	0,0259	0,0113	0,1741	0,2813	0,1591	-0,1152	0,2355	-0,3785	0,3107	-0,0513	-0,4312	0,0736	0,1207	0,3383	-0,2191	-0,0439	-0,2279	0,0971

CONTINUAÇÃO

Matriz de Correlação

47	0,3048	-0,0295	0,3881	0,4026	-0,0007	0,0802	-0,2538	0,0068	0,3569	0,3037	0,3585	-0,1778	0,0108	0,1132	0,6890	0,5083	0,5706	0,0530
48	-0,0167	0,2825	0,3936	0,2863	-0,0844	0,2144	-0,3045	0,0837	-0,1486	0,2476	-0,1224	-0,1789	-0,2482	0,2481	0,3044	0,1967	-0,3359	0,1035
49	0,0748	-0,2930	0,3219	0,0205	0,3897	-0,2127	0,1018	0,2176	0,3268	0,4993	0,2331	-0,3076	-0,1066	0,4590	0,5103	0,3780	0,2977	0,0848
50	0,2827	-0,1567	0,4268	0,1247	0,1622	0,2673	0,5222	-0,4563	-0,0137	-0,2923	-0,0579	0,1072	0,4284	0,2808	0,0564	0,0149	0,4139	-0,0022
51	-0,0637	0,3976	0,3824	0,4417	-0,1457	0,2171	-0,3690	-0,3430	-0,3940	-0,0222	-0,2563	-0,2906	0,3068	0,1988	0,0187	-0,2618	0,1284	-0,0436
52	-0,3276	0,1509	-0,1135	-0,4032	-0,1169	0,1748	0,1446	-0,0217	-0,4170	-0,2573	-0,1329	0,3526	-0,0753	0,1408	-0,2559	-0,2761	-0,2761	0,1052
53	0,1765	0,2666	0,3684	0,1977	-0,1347	0,5575	-0,1239	0,4163	-0,3020	-0,2520	-0,0365	0,5259	0,2068	-0,2453	0,1354	-0,2652	0,0366	-0,2107
54	0,3089	-0,2378	-0,1413	-0,3939	-0,1634	-0,1352	0,4992	-0,0107	0,0910	-0,3860	0,0679	0,0116	-0,2020	-0,3129	-0,3472	-0,2421	-0,1064	-0,3450
55	0,2091	0,5515	0,3142	0,0567	-0,1026	0,6525	0,0009	-0,2517	-0,5821	-0,5571	-0,3820	0,2451	0,3607	-0,0891	-0,3881	-0,3915	-0,1233	-0,4387
56	0,4860	0,1745	0,3055	-0,2751	0,5776	-0,0958	-0,0239	-0,1514	0,2521	0,4135	0,3479	-0,2354	0,1163	-0,0288	0,2531	-0,1555	0,2699	-0,1780
57	-0,2854	-0,0157	-0,1647	-0,0378	-0,1266	-0,0069	0,1770	0,1807	0,0009	-0,2948	-0,1873	0,1756	0,2866	0,3896	0,0705	0,4599	-0,0216	0,2325
58	0,2643	0,2780	0,2073	0,1213	0,2758	0,0604	-0,3966	0,3493	0,2548	0,4174	0,2053	-0,3007	0,0422	0,2781	0,6254	0,4242	0,2196	-0,0878
59	0,1129	-0,0723	-0,1810	-0,1448	-0,1382	0,2891	0,1359	-0,0618	-0,1502	-0,3518	0,0984	-0,0650	0,0280	0,2286	-0,0694	0,3237	-0,0111	-0,0585
60	0,2454	0,1359	-0,0105	-0,2116	0,4565	0,3509	-0,0673	0,0905	-0,1659	-0,0270	0,2028	-0,1580	0,2465	-0,1330	0,0402	0,1620	0,0548	-0,0609
61	-0,0805	-0,2690	-0,0752	0,2854	-0,0576	-0,0725	-0,2873	-0,3480	0,5180	0,5223	0,3067	-0,3229	-0,3841	0,2108	0,7604	0,7771	0,0162	0,2556
62	0,0273	0,2816	0,1574	-0,1500	0,0669	0,2396	0,1035	-0,3639	-0,5310	-0,3955	-0,2912	0,2676	0,5344	0,0675	-0,4347	-0,7222	-0,1054	-0,2571
63	0,4058	0,1324	0,2126	0,0750	0,2815	0,1897	-0,1848	0,1658	-0,0362	0,1300	0,1497	-0,0713	0,0261	0,0943	0,4554	0,4707	0,3138	0,0211
64	0,4490	0,4989	0,2294	-0,0624	-0,3429	0,0485	-0,1650	-0,0807	-0,2185	-0,2934	-0,1188	-0,0095	-0,2916	-0,0807	0,0116	0,1000	-0,0157	-0,0621
65	0,1322	0,3834	0,1294	0,2618	-0,3596	0,3979	-0,3764	-0,0807	-0,0531	-0,0422	0,1552	0,3694	-0,2916	-0,1864	0,4067	0,4226	-0,0157	0,2153
66	0,2141	0,1984	-0,1062	0,1329	-0,7664	0,2458	-0,2766	-0,0432	-0,4310	-0,3229	-0,0746	0,0263	-0,1772	-0,1350	-0,0792	-0,0790	-0,1769	-0,2819
67	0,1263	0,3206	0,0371	-0,2317	-0,0421	-0,1243	-0,0284	0,0001	0,2289	0,0066	-0,4916	-0,1832	-0,1899	-0,2806	-0,3986	-0,3612	-0,6107	-0,3532
68	-0,3484	0,0956	-0,0948	0,2889	0,2121	-0,0090	-0,3946	0,1450	0,0120	0,4744	0,0207	-0,1365	0,0952	0,1879	0,1064	-0,0375	-0,1369	0,0898
69	0,0861	0,3837	0,0182	0,0841	-0,4014	0,2160	-0,1799	-0,2182	-0,3026	0,0390	-0,0489	0,2166	0,1426	0,0767	-0,0886	-0,3204	-0,2112	-0,1862
70	0,4578	0,4634	0,2693	-0,0059	-0,0598	0,2144	-0,0779	-0,2286	-0,4286	-0,0312	-0,4203	-0,1553	0,3090	0,0569	-0,2512	-0,3156	-0,2248	-0,2743
71	0,0385	0,0856	0,3585	0,0925	-0,1623	-0,4319	-0,0314	0,0226	0,5226	0,2671	0,4989	0,1614	-0,1599	0,0725	0,3650	0,0879	0,3565	0,2873
72	0,1457	-0,3329	0,3589	0,1654	0,1923	-0,1942	0,2101	-0,3974	0,5226	0,2671	0,4989	0,1614	-0,1599	0,0725	0,3650	0,0879	0,3565	0,2873
73	-0,1463	-0,0192	-0,1384	-0,0645	0,3500	0,1544	-0,5205	0,3257	0,1985	0,3576	-0,0881	-0,3540	0,3078	0,1902	-0,0116	-0,0581	-0,0529	-0,2087
74	-0,0717	0,4268	0,3090	0,3072	0,1165	0,1544	-0,5205	0,3257	0,1985	0,3576	-0,0881	-0,3540	0,3078	0,1902	-0,0116	-0,0581	-0,0529	-0,2087
75	0,2741	0,1910	0,1516	-0,1235	0,2504	0,0122	0,0213	-0,0754	-0,2484	0,0001	-0,2182	-0,0668	0,5535	0,0839	-0,0148	-0,3462	-0,3780	-0,3274
76	0,0323	0,1540	-0,1675	-0,0958	-0,3126	0,4466	-0,0069	-0,1128	-0,2385	-0,1238	0,1120	0,1805	0,1617	0,4012	0,2441	0,1940	-0,0648	0,1713
77	-0,5130	-0,3395	-0,5739	-0,1746	-0,0028	-0,1617	-0,0243	0,2034	-0,0521	0,0300	-0,0064	-0,0322	0,1195	0,3528	-0,1149	0,0449	-0,0651	0,3278
78	-0,4786	-0,1137	-0,4373	-0,0609	-0,1336	-0,1309	-0,2185	-0,0461	-0,0977	0,0806	-0,0437	-0,1667	0,1880	0,2413	-0,0273	-0,0557	-0,0073	0,3306
79	0,3320	0,2522	0,0818	0,1855	-0,1776	0,6085	-0,0813	-0,2877	-0,3752	-0,0762	-0,3271	0,2349	0,1455	-0,1288	-0,2461	-0,2260	-0,4050	-0,2452
80	0,0981	0,0785	0,2314	0,1508	-0,2516	0,0748	0,1801	-0,1696	-0,0549	-0,1469	-0,0676	0,0283	0,2769	0,4995	0,2641	0,2327	0,1687	0,3221

CONTINUAÇÃO

Matriz de Correlação

47	0,2729	0,2396	-0,0580	-0,0329	0,3653	0,0808	0,4005	0,3284	0,3149	-0,0530	1,0000	1,0000	1,0000	1,0000	0,4666	1,0000	0,4757	-0,0048	1,0000
48	0,0428	0,0091	-0,1593	0,2959	-0,1828	0,0908	0,3857	0,1253	0,0928	-0,1453	0,1805	1,0000	1,0000	1,0000	0,4757	0,0281	-0,0191	1,0000	1,0000
49	0,3151	0,6328	0,2681	0,1121	0,0832	0,2607	0,5175	0,4314	0,2708	0,3319	0,4117	0,5634	1,0000	1,0000	0,4666	0,0281	-0,0191	1,0000	1,0000
50	-0,0195	-0,0058	0,2986	0,0570	0,0926	0,1413	0,2248	-0,0911	-0,2363	-0,0018	0,3825	-0,1721	0,1143	1,0000	0,4666	0,0281	-0,0191	1,0000	1,0000
51	0,0769	-0,2240	-0,3261	0,2438	0,4317	-0,1190	0,0338	-0,0419	-0,2715	-0,7106	0,3338	0,2051	-0,1322	0,2318	1,0000	0,4666	0,0281	-0,0191	1,0000
52	-0,3327	-0,1040	0,0612	0,2406	-0,0511	0,0347	-0,4324	-0,0770	-0,3087	-0,0391	-0,4583	0,0066	-0,0500	0,0584	0,0466	1,0000	0,4666	0,0281	-0,0191
53	-0,1168	-0,2887	-0,1575	-0,2288	0,1787	-0,1934	-0,0174	-0,3679	-0,3946	-0,4120	0,2583	-0,0955	-0,3881	0,3640	0,4757	-0,0048	1,0000	1,0000	1,0000
54	-0,3428	-0,0045	0,5510	-0,1391	-0,4127	-0,0727	-0,2411	-0,4931	-0,1720	0,5097	-0,4081	-0,3726	-0,1961	0,0378	-0,6188	0,0281	-0,0191	1,0000	1,0000
55	-0,1639	-0,3725	0,0888	0,1985	0,1571	-0,0618	-0,0715	-0,6260	-0,3252	-0,3371	-0,0224	-0,0261	-0,3805	0,4095	0,5296	0,1357	0,6613	0,0720	0,0805
56	0,3344	0,3942	0,0365	-0,0147	0,2054	-0,2836	0,0694	0,1612	-0,1341	0,2570	0,2859	-0,0570	0,2438	0,1810	-0,1287	-0,1570	0,0356	0,0805	0,0805
57	0,0335	-0,1483	-0,0071	0,1478	-0,0229	0,7012	0,2182	0,3430	0,3460	-0,1898	-0,0517	-0,0836	0,0233	0,1970	0,0255	0,0961	-0,2404	-0,1580	-0,1580
58	0,4041	0,2628	-0,2413	0,0208	0,3771	0,1558	0,3886	0,3430	0,4578	0,0707	0,6451	0,2013	0,3931	-0,0376	0,1541	-0,4099	-0,0453	-0,4208	-0,4208
59	0,0448	-0,1492	-0,0987	0,0827	-0,1735	0,4195	0,3629	0,1024	0,2869	0,3027	-0,0752	-0,0845	0,0566	0,2034	-0,5033	0,0305	-0,0172	0,2746	0,2746
60	0,2973	-0,0576	-0,3601	0,0345	0,1006	0,2747	0,4791	0,1894	0,1732	0,1767	0,0981	-0,0065	0,0935	0,1680	-0,3086	-0,1298	0,2006	0,0473	0,0473
61	0,2552	0,2892	0,0318	-0,0534	-0,2598	0,3464	0,4566	0,6636	0,6136	0,1788	0,4127	0,3735	0,5366	-0,2709	-0,2919	-0,3041	-0,4428	-0,3099	-0,3099
62	-0,2155	-0,3545	-0,2332	0,0567	0,1787	-0,3521	-0,4269	-0,5955	-0,5955	-0,0428	-0,3057	-0,2930	-0,4785	0,3311	0,3116	0,3836	0,4612	0,1719	0,1719
63	0,2907	0,1148	-0,1377	0,4674	0,1665	0,4844	0,6830	0,4677	0,4573	0,0782	0,5895	0,2046	0,4559	0,3578	0,0966	-0,2284	-0,0720	-0,2676	-0,2676
64	-0,1316	-0,3225	-0,2811	0,5124	0,1483	0,1331	0,1368	0,0606	0,0092	-0,2661	0,2658	0,0979	-0,1827	0,2791	0,3620	-0,0671	0,1839	-0,0141	-0,0141
65	-0,0080	-0,3416	-0,2263	0,2457	-0,1592	0,3975	0,2588	0,2344	0,2518	-0,4540	0,4370	0,1986	-0,1408	0,1328	0,2538	0,2538	-0,2014	0,3378	-0,2095
66	-0,3084	-0,4619	-0,2367	0,4795	-0,0067	-0,1480	-0,0735	-0,2810	-0,0721	0,0760	0,1344	0,1929	-0,2367	-0,0024	0,3183	0,1069	0,1798	0,1745	0,1745
67	-0,2779	0,0203	0,0142	0,2336	-0,1946	-0,3778	0,2943	-0,3417	-0,2429	0,0268	-0,5072	0,1850	-0,2043	-0,4099	-0,0027	-0,1218	-0,2029	0,1610	0,1610
68	0,3922	-0,0176	-0,5046	-0,2535	0,1341	-0,1216	0,1423	0,1305	0,0611	-0,2142	-0,0590	0,3890	0,1033	-0,5885	0,0674	-0,3495	-0,1165	-0,4661	-0,4661
69	-0,1354	-0,4482	-0,5027	-0,0480	0,3299	-0,3476	-0,3685	-0,2081	-0,4830	-0,3788	0,0320	0,0765	-0,3413	-0,1950	0,3804	0,1379	0,3422	-0,1847	-0,1847
70	-0,1959	-0,3084	-0,4062	0,1895	0,4688	-0,2675	-0,0952	-0,1806	-0,4064	-0,3855	0,1174	-0,0298	-0,3132	0,0529	0,5838	0,0261	0,3776	-0,2833	-0,2833
71	-0,1764	0,2797	0,3425	-0,1401	0,0013	-0,2953	-0,2662	-0,2218	-0,1063	0,0872	0,0699	0,0430	0,0593	-0,1221	0,2163	-0,1484	-0,0569	0,1967	0,1967
72	0,1692	0,2515	0,3995	-0,3502	-0,2411	-0,0124	0,1151	0,4261	-0,1992	0,0535	0,2530	0,0152	0,3697	0,2992	-0,2628	-0,0357	0,0052	0,1788	0,1788
73	0,3340	0,3847	-0,1254	-0,2055	0,3071	-0,2203	0,0013	0,1138	0,1912	0,1418	-0,1091	-0,0634	0,1050	-0,4733	0,0084	-0,2097	-0,3129	-0,2413	-0,2413
74	0,1631	-0,0411	-0,4860	0,0655	0,3769	-0,5176	-0,0835	0,0106	-0,2514	-0,4186	0,2128	0,3773	-0,0371	-0,1364	0,6835	0,0066	0,2530	-0,7123	-0,7123
75	0,0749	0,0075	-0,3096	-0,0953	0,7051	-0,3094	-0,0514	-0,3270	-0,2805	-0,1583	0,1395	-0,2166	-0,2019	0,0808	0,4386	-0,1910	0,4235	-0,1693	-0,1693
76	0,0080	-0,3660	-0,2306	0,4359	-0,0642	0,3143	0,1555	0,1069	0,0434	-0,1030	-0,1779	0,1066	-0,0724	0,3651	0,1327	0,2336	0,1109	-0,1548	-0,1548
77	0,1703	0,0268	-0,0028	0,2105	-0,1492	0,2577	0,1058	0,1778	0,2876	0,0673	-0,3894	-0,0744	-0,0235	-0,1950	-0,1069	0,1366	-0,4151	-0,1343	-0,1343
78	0,0982	-0,0627	-0,2935	0,3130	0,0800	-0,1568	-0,1740	0,1994	0,0785	-0,0862	-0,1345	-0,0995	-0,2407	-0,0285	0,2427	0,2286	-0,2542	-0,4430	-0,4430
79	-0,1224	-0,4505	-0,2494	-0,0354	-0,0502	-0,2134	0,1034	-0,4408	-0,2957	-0,3535	0,0062	0,0773	-0,4509	-0,0051	0,3242	-0,2545	0,5888	-0,1071	-0,1071
80	-0,1386	-0,2409	-0,0986	0,2735	0,2424	0,2394	0,0598	0,2683	-0,0682	-0,2157	0,3983	0,0960	0,1579	0,5337	0,3941	0,2551	-0,0117	-0,3527	-0,3527

CONTINUAÇÃO

Matriz de Correlação

	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72		
47																				
48	0,0710																			
49	-0,0328	1,0000																		
50	0,0188	0,4612	1,0000																	
51	-0,0339	0,0551	0,1385	1,0000																
52	0,1054	0,4307	-0,0706	-0,1396	1,0000															
53	-0,7366	-0,0840	0,1995	0,3807	0,1271	1,0000														
54	0,6142	0,1428	-0,1888	-0,2849	0,0393	0,0463	1,0000													
55	0,0710	1,0000																		
56	-0,0328	-0,5377	0,0392																	
57	0,0188	0,4612	0,1385	1,0000																
58	-0,0339	0,0551	0,1385	-0,1396	1,0000															
59	0,1054	0,4307	-0,0706	0,2382	0,7979	1,0000														
60	-0,7366	-0,0840	0,1995	0,3807	0,1271	0,0463	1,0000													
61	0,6142	0,1428	-0,1888	-0,2849	0,0393	0,0463	1,0000													
62	0,0710	1,0000																		
63	0,0152	0,4559	0,0780	0,5379	0,4311	0,5478	-0,7712	1,0000												
64	0,4071	0,0714	0,2863	0,1696	-0,0315	-0,0001	-0,1545	0,2007	1,0000											
65	0,1076	0,0093	0,1665	0,1708	0,1171	0,1987	0,2969	-0,1635	0,4612	1,0000										
66	0,2505	-0,2940	-0,1134	-0,1643	-0,0057	-0,1942	-0,0813	0,2822	0,0444	0,5040	1,0000									
67	-0,1315	0,0517	-0,2994	-0,2480	-0,3289	-0,3614	-0,2647	0,0935	-0,3608	0,1223	-0,2543	1,0000								
68	-0,2592	-0,1145	-0,1313	-0,2638	-0,1813	0,0947	0,2830	-0,2309	-0,1926	-0,4335	-0,1155	-0,2210	1,0000							
69	0,1651	-0,1449	-0,1714	0,0037	-0,3456	-0,2154	-0,1742	0,2891	-0,3446	0,2022	0,0865	0,5129	0,0281	1,0000						
70	0,4288	0,0909	-0,1533	0,1764	-0,4226	-0,2043	-0,3653	0,3272	-0,0565	0,4220	0,0225	0,3032	0,3822	0,3881	1,0000					
71	0,1085	0,0311	-0,0475	0,1476	-0,7137	-0,6620	-0,1235	0,0253	-0,3059	0,2456	-0,1428	0,0865	0,3817	-0,1127	0,2635	1,0000				
72	-0,3940	0,3451	-0,1975	-0,1389	0,2514	0,1898	0,2995	-0,1860	0,2409	-0,1617	0,1695	-0,2546	0,3817	-0,1127	0,0951	0,1879	1,0000			
73	-0,1102	0,0516	0,0328	0,4675	-0,3674	-0,1694	0,0576	-0,1633	-0,2232	-0,2793	-0,5893	-0,3798	0,2416	-0,2213	0,4719	0,3882	0,2635	1,0000		
74	0,2611	0,1830	-0,4122	0,2963	-0,4997	-0,1645	0,1985	0,1985	-0,0310	0,0197	-0,0498	0,1048	0,2889	0,4370	0,4370	0,4370	0,4370	0,4370	1,0000	
75	0,5003	0,1400	-0,1365	-0,3317	-0,3935	-0,0482	-0,5284	0,3267	-0,1812	0,1132	-0,3237	-0,0483	0,1138	0,1700	0,3786	0,4055	0,4055	0,4055	0,4055	1,0000
76	0,1626	-0,0839	0,4059	0,1341	0,1347	0,1895	0,0905	0,2125	0,2876	0,4514	0,5038	0,3785	-0,3717	-0,1835	0,2219	0,0077	0,0077	0,0077	0,0077	1,0000
77	-0,2218	-0,4406	0,5025	-0,0766	-0,0744	-0,1243	0,1058	-0,0859	-0,1681	-0,2130	-0,1924	-0,1798	-0,1740	0,2107	-0,3390	-0,4171	-0,0977	-0,0977	-0,0977	1,0000
78	-0,1336	-0,2788	0,2587	-0,1147	-0,2518	-0,2145	-0,0081	0,1813	-0,1313	0,0760	-0,0364	0,1533	0,0211	0,0676	-0,0243	-0,0243	-0,0690	-0,0690	-0,0690	1,0000
79	0,4784	-0,0941	-0,2065	0,0779	-0,2380	0,0169	-0,2517	0,1734	-0,2244	0,0849	0,1781	0,1750	0,1877	0,2763	0,4446	0,5665	-0,0693	-0,0693	-0,0693	1,0000
80	0,0494	-0,1785	0,5003	0,1392	-0,1410	-0,2324	0,0906	0,0722	0,2428	0,5108	0,2289	0,3142	-0,3388	-0,3499	0,2936	0,3017	0,0290	0,0290	0,0290	1,0000

Matriz de Correlação

CONTINUAÇÃO

	73	74	75	76	77	78	79	80
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								
61								
62								
63								
64								
65								
66								
67								
68								
69								
70								
71								
72								
73	1,0000							
74	0,3273	1,0000						
75	0,5044	0,4523	1,0000					
76	-0,4346	-0,1048	-0,1791	1,0000				
77	0,3086	-0,1768	-0,1421	0,2965	1,0000			
78	0,1260	0,2887	-0,0580	0,4237	0,6532	1,0000		
79	0,0435	0,3842	0,5075	0,1408	-0,1153	-0,0747	1,0000	
80	-0,2851	0,0503	0,0101	0,6005	0,0000	0,2901	-0,1220	1,0000

ANEXO 05 – MATRIZ DE CARREGAMENTOS FATORIAIS

Continuação

Matriz de Carregamentos Fatoriais

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0,0668	0,0206	0,4201	0,2416	-0,2047	0,3061	0,2357	0,0353	0,0415	0,1236	0,1437	-0,1358	-0,0890	-0,2275	-0,4960	0,3048	-0,2775
2	-0,4046	0,1550	-0,1916	-0,1905	0,1194	0,1109	0,0488	0,2426	-0,0163	0,0353	0,1022	-0,0270	-0,0578	0,2773	-0,6407	-0,2039	0,3059
3	-0,0020	-0,0315	0,2008	-0,0019	0,0287	-0,2194	0,0571	-0,1170	-0,2423	-0,1532	-0,1362	0,4646	0,3329	-0,2166	0,5430	0,2329	0,2172
4	-0,2641	0,0210	0,0724	-0,0222	0,0018	-0,1452	-0,0400	0,2174	-0,0555	0,0602	0,1469	0,1234	-0,3021	0,0306	0,3638	-0,5031	-0,5712
5	-0,3038	0,0906	0,0413	-0,2308	-0,1446	-0,0854	0,3459	0,3502	-0,0530	0,1650	0,0957	0,0025	-0,1654	0,1155	-0,2755	-0,5949	0,2453
6	0,3926	-0,3381	0,0667	-0,0173	0,0290	-0,1738	0,1194	-0,0894	-0,0160	-0,1087	-0,0145	-0,2270	-0,0169	-0,0428	0,2175	0,3275	-0,6341
7	-0,3704	0,0591	-0,0167	-0,1623	0,0878	0,1334	0,2000	-0,0212	-0,1169	-0,0977	0,0616	0,0772	0,3708	-0,0717	0,0145	-0,2048	
8	-0,0281	0,1499	-0,4620	0,1535	0,0598	-0,1158	-0,1032	-0,0279	0,0530	0,2388	-0,0309	-0,0123	-0,0185	-0,1593	0,0573	0,1917	-0,6444
9	0,1261	-0,3008	0,0994	0,1781	-0,2069	-0,3081	0,0514	-0,1090	0,0524	-0,0864	0,1469	0,2498	-0,0185	-0,1037	0,2262	0,4464	-0,5707
10	-0,2575	-0,2554	-0,2153	-0,0759	-0,0220	-0,1822	0,2104	-0,1977	-0,0874	0,0799	-0,0204	0,0550	0,3536	-0,2994	-0,0512	-0,5286	-0,4084
11	-0,1241	-0,5670	-0,1755	-0,2168	0,1206	0,0473	-0,1815	-0,0994	-0,0461	-0,1391	-0,0702	-0,1006	0,1165	0,0883	-0,4301	-0,0386	-0,5002
12	-0,1964	-0,4679	0,2862	-0,0752	-0,1675	0,1080	-0,2160	-0,1394	0,0087	0,1585	-0,0306	0,2706	-0,0878	-0,3985	-0,1503	-0,0822	-0,4881
13	0,1421	-0,1615	0,0045	0,0688	-0,1991	-0,1689	-0,0425	0,2394	-0,0311	0,1612	0,0032	-0,0906	0,4627	-0,7425	-0,0324	-0,0693	-0,1015
14	0,1494	0,1407	0,3263	0,4890	0,1190	0,0884	-0,0453	-0,1998	-0,1693	0,0121	0,0598	-0,4946	-0,0995	-0,0738	-0,1642	-0,2417	-0,3806
15	0,0678	0,1339	0,0603	-0,1379	0,0114	0,1613	0,4818	-0,2651	0,0971	0,1385	0,1332	0,0638	0,0026	-0,3979	-0,0343	0,5755	-0,2695
16	-0,0596	0,1364	-0,3012	-0,1252	-0,5104	-0,0365	0,1327	-0,1693	-0,3192	-0,1209	0,0786	-0,1015	-0,1708	0,1291	-0,0201	-0,0117	-0,6146
17	0,4797	0,1109	-0,1743	-0,3935	0,4025	0,0369	0,0619	-0,0704	-0,0793	0,1899	0,2671	0,2041	0,3021	0,0236	0,0591	0,3416	-0,1721
18	0,1158	-0,2354	0,3677	-0,2743	-0,1976	0,1262	-0,0059	0,0636	0,0323	0,1916	-0,0230	-0,5977	0,0137	0,1881	-0,3771	0,0956	-0,2033
19	0,1196	-0,3130	0,0077	0,2164	0,1976	0,0032	-0,1413	-0,0314	0,0887	0,1215	0,0711	0,3036	-0,3011	-0,6880	-0,0843	-0,1842	-0,0098
20	0,2612	-0,3582	-0,1280	-0,3313	0,2059	0,2274	-0,1691	-0,0718	0,0219	-0,3442	-0,0015	-0,0319	-0,0705	0,0967	-0,0202	-0,4512	0,4298
21	0,4996	0,0592	-0,0069	-0,2461	0,0050	-0,0955	-0,1166	0,0134	0,0253	-0,2636	-0,0954	-0,1662	0,3143	-0,5501	-0,2972	-0,2176	-0,0155
22	0,6598	0,2405	-0,1305	0,0259	0,0449	0,1140	0,1683	-0,2279	-0,1718	0,2164	0,0940	-0,4020	0,2990	0,0428	-0,1055	-0,1786	-0,0836
23	-0,5400	-0,0517	-0,0686	-0,1047	0,0621	0,0409	-0,0265	0,0468	0,2750	-0,2052	0,0508	-0,1662	0,2253	-0,4833	-0,2657	-0,1033	-0,3752
24	0,1617	-0,2201	-0,3343	0,1905	0,0108	0,0794	-0,2178	0,2197	0,0727	0,2244	0,2170	-0,2571	0,1910	-0,1991	0,3676	-0,4146	0,3276
25	-0,2631	0,0919	0,3876	0,2017	-0,3031	0,1883	-0,0668	0,1377	0,1504	-0,1000	-0,0380	-0,2285	0,1196	-0,5175	0,1015	0,4454	0,0476
26	0,0011	-0,6421	0,2286	-0,0918	0,1563	-0,0396	-0,1591	0,2284	-0,0790	0,0549	-0,0080	-0,2241	-0,1986	0,4056	0,1067	0,0404	-0,3560
27	0,2366	0,3957	0,0213	0,1745	0,1315	-0,0430	-0,1135	-0,1232	0,0051	-0,1633	0,0822	0,0870	-0,1510	0,0037	-0,2350	0,3673	-0,6614
28	-0,0376	0,2651	-0,2461	0,1136	0,1290	0,0612	-0,0433	0,1798	0,1118	-0,0437	0,0615	0,4449	0,0656	0,2064	-0,5008	-0,1364	-0,4596
29	-0,16916	0,4958	-0,2776	-0,0483	0,0851	-0,0010	-0,1501	0,0128	-0,1214	0,0867	-0,2100	0,3007	-0,4585	0,0548	-0,0014	-0,0031	-0,4415
30	-0,0916	0,3727	-0,3437	0,0444	0,0470	-0,0501	0,1360	0,3688	-0,1157	-0,2911	0,2098	-0,1266	0,0964	-0,0421	0,3716	-0,1935	0,3931
31	-0,4392	0,1879	0,2459	0,2093	0,0780	-0,0253	-0,0770	-0,3969	-0,2034	-0,1430	0,1039	-0,1406	0,2486	-0,2461	-0,3312	-0,3696	0,1905
32	-0,0812	0,3521	0,3509	0,1374	-0,0656	0,2534	-0,3321	0,0374	-0,2086	-0,0325	0,0947	0,1043	0,5916	0,0084	-0,1122	-0,2533	-0,1834
33	0,4779	0,2096	-0,0173	0,0062	0,1459	0,0755	-0,3624	0,0741	-0,0198	0,0184	-0,0198	0,0764	-0,0084	0,0683	-0,0001	-0,2749	-0,6589
34	0,3193	-0,1681	0,1591	0,1866	0,0332	-0,0517	0,0150	0,0940	0,0234	-0,0288	0,0786	-0,1844	-0,1530	0,1687	0,3821	-0,2112	-0,6754
35	-0,1890	0,5606	0,2289	-0,0428	0,1919	0,0818	0,1179	0,0613	-0,0234	0,2577	-0,2197	-0,1739	-0,3782	-0,1926	-0,2445	-0,1113	-0,3677
36	0,0698	0,2636	0,1803	-0,1687	-0,2475	-0,2124	-0,1346	-0,2134	0,5445	-0,1638	-0,2576	-0,2420	0,0570	0,3397	0,2310	-0,0461	-0,2507
37	-0,3533	0,3536	-0,2253	0,1704	0,2104	0,2186	0,1823	-0,0887	-0,1751	0,0146	-0,0601	0,2349	0,0565	0,0922	-0,3206	-0,3257	-0,4139
38	-0,1993	0,1755	0,0808	-0,0367	0,2145	-0,1525	0,0401	0,3613	-0,0287	0,0180	-0,0182	0,3059	0,0454	-0,0896	-0,4074	0,2057	-0,5163
39	0,1557	0,2381	0,1633	0,0654	0,1960	0,0108	0,0851	0,3503	0,1875	0,0210	0,0686	0,0856	0,1862	0,4002	-0,1198	0,6446	-0,2118
40	-0,1445	-0,3100	0,2899	-0,1964	-0,0418	0,1776	0,2871	0,1037	0,1387	0,1994	-0,1172	0,4978	0,0483	0,0109	0,3588	-0,4019	0,1332
41	-0,3325	0,0979	0,2993	-0,0764	-0,0421	-0,0869	0,0459	0,1020	-0,0986	0,0074	-0,1152	-0,1217	-0,4063	0,0501	-0,5122	-0,5082	0,0871
42	-0,0837	-0,1715	0,2254	0,0866	0,1015	0,2167	0,1474	0,2823	0,0012	-0,2486	-0,0592	-0,3296	0,0835	0,0547	0,6049	-0,2819	-0,3185

Continuação

Matriz de Carregamentos Fatoriais

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
43	-0,0371	-0,2207	-0,0717	0,2897	-0,0223	-0,0689	0,2880	-0,0095	-0,0322	0,0337	-0,0910	0,0315	0,3223	-0,2860	0,1350	-0,4984	-0,5117
44	-0,1033	0,2934	-0,0114	-0,1363	-0,1532	-0,1642	0,1199	-0,1246	0,0696	-0,2988	0,1529	0,1014	0,0104	0,2483	0,0558	-0,3792	-0,6611
45	0,1293	-0,4064	0,1736	0,0622	0,3186	-0,0603	0,0849	0,0396	0,0408	-0,1317	0,0255	-0,1529	-0,1706	0,3439	0,2805	-0,1229	-0,6202
46	-0,2041	-0,3987	0,1985	0,0275	-0,0083	-0,1197	-0,3625	-0,0044	-0,0712	0,2708	0,1288	0,1362	-0,0843	-0,2565	-0,0076	0,4580	-0,4479
47	0,4392	0,2271	0,1112	0,1155	0,0864	-0,0867	-0,0356	0,0781	0,0124	0,2104	-0,0689	-0,0412	-0,3037	-0,2390	-0,1661	-0,5396	-0,4323
48	0,4146	-0,2564	-0,1691	-0,1610	-0,2357	0,0769	-0,0878	0,1335	-0,0640	0,0417	-0,4149	0,2269	0,4879	0,0927	-0,0021	-0,3121	-0,1822
49	0,0334	-0,0146	0,1639	-0,1457	-0,1634	0,0309	-0,0816	0,3606	-0,1169	0,0285	-0,2622	0,1404	0,2912	-0,1709	-0,1485	-0,1432	-0,7256
50	-0,0296	0,3688	0,3315	0,0437	0,0618	-0,0108	-0,0558	-0,0428	0,1504	0,1485	0,0200	-0,1920	0,1107	0,7332	-0,1796	-0,2295	0,0350
51	0,2861	0,2850	0,1431	-0,1449	0,1354	-0,0584	0,2369	0,0696	0,0034	0,1172	-0,0568	-0,0276	0,1717	0,0568	-0,2524	-0,6101	0,4710
52	-0,2884	0,1092	0,1296	-0,4911	-0,1692	-0,1739	-0,3076	0,3593	-0,0242	-0,0515	0,1367	0,0929	0,1917	0,1080	0,3327	0,0540	0,3726
53	0,1600	0,2043	-0,3840	0,1709	0,1789	-0,3080	-0,0828	0,1154	-0,0307	0,0509	-0,0833	-0,2342	-0,1331	-0,3690	-0,0167	-0,3235	0,5212
54	-0,1239	-0,1716	-0,0482	0,0971	0,1047	0,0618	-0,0798	-0,0136	-0,2060	-0,0999	-0,0758	0,1221	-0,2277	-0,4473	0,2605	-0,7007	0,1205
55	0,0324	-0,1554	0,0200	-0,0087	0,4226	0,0163	-0,0025	0,0359	-0,0548	0,0485	-0,1033	-0,1205	0,0925	-0,4037	-0,0143	-0,2993	0,6821
56	-0,1557	-0,0488	-0,2223	-0,2373	0,1841	0,1720	-0,2964	-0,0234	-0,0264	0,1094	0,1101	0,2822	-0,2415	-0,4073	-0,4070	-0,1473	-0,2637
57	-0,0965	0,1083	0,5377	0,2570	0,1257	0,0442	0,0930	-0,0696	-0,1513	-0,3494	0,0219	-0,2135	0,2224	0,2419	0,4790	-0,1863	-0,1086
58	0,1566	-0,2109	0,1037	0,1803	0,3067	0,0708	-0,2342	0,0763	-0,0038	-0,0927	0,0837	0,0622	-0,1369	0,0321	-0,3692	-0,5694	-0,4378
59	-0,3830	-0,2693	-0,2640	-0,0935	-0,0764	0,0253	0,0154	-0,2018	-0,2029	0,0857	0,0492	-0,2721	-0,1048	-0,3086	0,5367	-0,0717	-0,3126
60	-0,4697	-0,2694	-0,0970	0,0094	0,0609	0,0379	-0,1650	0,0756	-0,0731	-0,0023	0,1588	-0,0095	0,0865	-0,3242	0,3137	-0,1756	-0,1229
61	0,3218	0,0068	-0,0970	0,1576	-0,1923	-0,0564	-0,0684	0,0181	0,0731	-0,0943	0,1266	0,1068	0,0439	0,3137	0,1756	-0,3861	-0,7905
62	-0,3168	0,0642	0,0023	-0,2426	0,1262	-0,0564	-0,2839	-0,2589	-0,0943	0,1162	0,1266	-0,0095	0,0865	-0,3242	0,3137	-0,1756	0,6750
63	-0,1334	-0,1431	0,0398	-0,2009	0,1372	0,1650	0,1400	-0,0016	0,1077	0,0944	0,0667	0,0632	-0,0997	-0,3468	0,1962	-0,6271	-0,5189
64	0,2817	-0,1014	0,3347	-0,0419	0,1313	-0,0091	0,1014	-0,3048	-0,0164	-0,2774	-0,1152	0,3629	-0,2277	-0,2314	0,2674	-0,4449	0,2568
65	0,4398	0,1443	-0,2569	0,0534	0,1714	0,1664	0,0647	-0,0324	0,2181	0,1250	-0,0724	0,0677	-0,2400	0,0026	0,5100	-0,4759	-0,0246
66	0,3463	-0,1121	0,1042	-0,0749	-0,1879	-0,0003	0,0318	-0,1171	-0,3859	-0,0471	0,3344	0,3629	-0,2277	-0,2314	0,2674	-0,4449	0,2568
67	0,1385	-0,4750	-0,0138	-0,1632	-0,0811	0,0962	0,2072	0,0307	0,1146	0,1283	0,1283	0,4133	0,1590	0,0794	0,3047	-0,2067	0,4071
68	-0,0301	-0,1242	-0,3979	0,2260	-0,1520	0,2088	-0,0684	-0,1014	-0,0869	0,0239	-0,1938	-0,1491	0,2496	0,5519	-0,4503	-0,1712	-0,1260
69	0,2088	0,0975	-0,0687	0,1193	-0,4405	0,1301	-0,2662	0,0171	-0,2819	-0,0486	0,0655	0,0892	-0,2555	0,1928	-0,2124	-0,2796	0,5684
70	0,1705	-0,1527	0,1630	0,0816	-0,2615	-0,0587	0,1182	0,0763	0,0605	-0,1704	0,2524	0,0976	-0,1614	-0,1137	-0,1105	-0,3926	0,5817
71	0,4650	0,0369	0,3768	-0,0111	0,3436	-0,0348	-0,0679	0,0777	0,1641	-0,2208	-0,0948	0,0467	-0,0450	-0,4425	0,1149	-0,3936	0,2983
72	0,1050	0,5237	-0,3128	-0,2592	-0,2289	0,0010	-0,0222	-0,0305	-0,0032	0,0131	-0,1908	-0,0286	-0,0450	-0,4425	0,1149	-0,3936	0,2983
73	-0,1924	-0,2765	0,1632	0,1888	0,2276	-0,1699	0,0418	-0,0032	-0,1644	-0,1318	-0,1716	0,0391	0,1558	0,3937	-0,6607	0,0326	-0,1450
74	0,2158	-0,0143	-0,1463	-0,2125	-0,0638	-0,1072	-0,0233	-0,0417	0,1875	0,2083	0,0508	0,0914	0,1912	0,2179	-0,6116	-0,4651	0,2852
75	-0,1648	-0,1437	0,2062	0,3157	0,0915	-0,1872	0,0149	0,0796	-0,0205	-0,0454	-0,0768	-0,1262	-0,1576	-0,1182	-0,5650	-0,2871	0,4276
76	-0,0406	0,1688	0,1908	0,2341	-0,0055	0,1751	-0,0221	-0,1051	0,1050	-0,0003	0,0003	0,2437	0,0347	0,0073	0,5665	-0,4813	0,1216
77	-0,4168	0,0591	0,2535	0,2620	0,2111	-0,0822	-0,0021	-0,1079	0,1175	0,1315	-0,0913	0,0904	0,3976	0,5430	0,2542	0,0810	-0,0877
78	-0,1991	0,2266	0,3359	-0,0141	0,0542	-0,1091	-0,0630	-0,3220	0,2887	0,3562	0,1401	0,2472	0,1794	0,5304	0,0929	-0,1269	0,1435
79	0,2081	-0,1978	-0,3248	0,6041	-0,0668	-0,0290	0,0249	0,0196	0,1939	0,0751	0,0627	-0,0345	0,0876	-0,0816	-0,1509	-0,2425	0,5379
80	0,2145	0,3389	0,6446	-0,0352	-0,3038	-0,0164	-0,1793	-0,0665	-0,0009	-0,0644	0,0472	0,0110	0,0194	-0,0877	0,2352	-0,4582	0,0488

**ANEXO 06 – MATRIZ DE CARREGAMENTOS FATORIAIS
ROTACIONADOS**

Matriz de Carregamentos Fatoriais Rotacionados

Continuação

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0,0204	-0,1320	-0,3198	0,1734	0,1807	0,1462	0,3143	-0,0145	0,1621	0,1332	0,4114	0,1203	0,1042	0,4607	-0,4032	-0,1491	0,1584
2	-0,0233	-0,0928	0,3598	-0,3449	-0,0957	-0,0327	-0,5333	0,0599	-0,1572	-0,0199	-0,1739	-0,0777	0,2792	-0,5134	-0,0574	-0,1584	0,0107
3	-0,1641	-0,0974	-0,0694	-0,1371	-0,0713	-0,2279	-0,0291	-0,2157	-0,2613	0,1477	0,1470	-0,0699	0,1631	-0,1788	0,6789	0,1633	0,0472
4	0,6130	0,0594	0,4263	-0,0085	0,2262	0,1494	0,0363	0,3920	-0,0050	-0,0050	0,1341	0,0564	0,0216	0,2095	-0,3291	0,1348	0,0391
5	-0,0912	0,5669	0,1380	0,0370	0,0212	0,5244	-0,2591	0,1068	0,0014	-0,0743	0,1090	0,1330	0,3588	-0,2461	-0,1177	-0,0406	0,2278
6	0,6018	-0,3022	-0,0071	0,1170	0,1921	-0,3262	0,2089	-0,2013	0,3243	-0,2485	0,1268	0,2612	-0,1685	0,2178	0,0852	0,0920	-0,1232
7	0,0545	0,3449	0,4882	-0,0563	0,0587	0,0885	-0,1959	0,0805	0,3213	0,4063	-0,0828	0,1433	0,3737	-0,0189	0,1964	0,2081	0,2051
8	0,1324	-0,0397	0,3462	-0,0394	0,1692	-0,5249	0,1188	0,2953	0,2706	-0,0359	0,3331	-0,1733	-0,3180	0,1192	-0,0136	0,2527	-0,0429
9	0,6527	0,1641	0,3070	0,1696	0,3015	0,0705	0,2663	-0,0743	-0,0297	0,1476	0,1073	-0,1447	0,1011	0,0529	0,4258	0,0553	-0,0749
10	0,1133	0,3221	0,6730	0,0886	0,0658	-0,0058	0,1567	-0,0241	0,2824	0,0784	0,1318	0,0846	0,1356	0,1529	0,3337	0,2987	0,1488
11	0,2011	0,0994	0,1900	0,0310	0,1232	-0,0196	0,4961	-0,0945	0,5154	-0,2061	-0,3423	0,1824	-0,1330	0,0794	-0,0934	0,3013	0,1698
12	0,1630	-0,1197	0,2367	0,2128	0,0814	0,1010	0,8327	0,0539	0,1101	0,1432	0,0573	0,0372	0,2637	0,0271	0,0490	0,1091	0,0850
13	-0,0014	-0,0147	0,2272	0,4981	-0,2600	-0,1051	0,1555	-0,2321	0,0929	0,0929	0,5684	-0,0567	-0,0702	-0,0149	0,0652	0,3731	0,0886
14	0,2409	0,2034	0,0820	0,1876	0,0175	-0,0252	0,1030	0,1099	0,0755	0,1826	-0,0459	0,8531	-0,1187	0,0375	-0,0914	-0,0774	-0,1030
15	-0,2055	-0,2821	-0,0732	0,1565	0,1918	-0,4682	0,1542	0,0161	0,0920	-0,1777	0,4176	0,2200	0,2572	0,4048	0,0033	-0,2503	0,0182
16	0,1191	-0,0238	0,3481	0,1315	0,8391	0,0014	-0,0003	0,1529	-0,0325	-0,0896	0,0190	0,2191	-0,1861	0,0898	0,0653	0,0923	0,0141
17	0,0210	0,1501	-0,2493	0,0348	-0,0099	-0,9091	0,0089	0,0026	-0,0273	-0,1059	-0,0016	0,1784	-0,0108	-0,0149	0,0714	0,0087	0,1861
18	0,1650	0,1322	-0,1236	-0,0071	0,0715	0,1626	0,2757	-0,0515	-0,0599	-0,0214	0,1182	0,5464	0,0753	0,0791	-0,6176	0,0851	-0,0897
19	0,0542	-0,0968	0,1265	0,6740	0,0488	0,2385	0,2757	-0,0515	-0,0599	-0,0214	0,1182	0,5464	0,0753	0,0791	-0,6176	0,0851	-0,0897
20	0,0081	0,2383	-0,0063	0,2660	-0,3047	0,0697	-0,1377	-0,2895	-0,0424	-0,0078	-0,7788	0,0511	0,1425	-0,0926	0,0342	0,0645	0,0116
21	-0,0569	0,2042	-0,1188	0,7008	-0,2204	-0,1395	-0,0312	-0,1362	0,1977	0,2328	0,0278	0,2510	-0,1621	0,1135	0,1443	0,3315	-0,1300
22	0,1248	0,3869	-0,0210	0,1372	0,0312	-0,3893	0,2940	-0,0987	-0,2361	0,1913	0,0504	0,2726	-0,2796	0,4977	-0,2048	0,1545	0,0089
23	-0,1166	0,0813	0,2696	0,1092	0,0530	0,1458	0,2144	0,0805	0,1614	0,2998	0,4793	-0,0128	0,0615	0,0023	-0,2133	-0,0741	-0,1724
24	0,1634	0,1589	0,4508	0,2015	-0,4743	0,1479	0,1045	-0,1796	0,1614	0,2998	0,4793	-0,0128	0,0615	0,0023	-0,2133	-0,0741	-0,1724
25	-0,2307	-0,6473	-0,0492	0,1484	-0,0814	0,1479	0,1045	-0,1796	0,1614	0,2998	0,4793	-0,0128	0,0615	0,0023	-0,2133	-0,0741	-0,1724
26	0,6717	-0,1246	-0,0031	-0,2553	-0,0408	0,1160	0,3366	-0,1837	0,0320	0,2730	-0,2192	0,2079	-0,0270	-0,2140	-0,2230	0,0816	0,1205
27	0,2333	-0,1544	-0,3427	0,1003	0,4027	-0,2996	0,1716	0,4524	0,2465	0,0740	0,1059	0,0566	-0,3365	0,2627	0,0425	0,1915	0,0668
28	0,0521	0,3787	-0,1689	0,0423	0,5276	-0,0348	0,1301	0,8851	0,3750	0,0669	0,0869	0,0869	0,0011	0,0364	0,0684	-0,0346	-0,0304
29	0,0347	-0,0926	0,0647	-0,0064	0,3270	-0,0337	0,0171	0,8851	0,3750	0,0669	0,0869	0,0869	0,0011	0,0364	0,0684	-0,0346	-0,0304
30	-0,1079	-0,0418	0,1582	0,1148	-0,0659	0,0003	-0,8324	-0,0317	-0,0055	0,0895	0,0174	-0,1772	-0,0523	-0,3238	0,0777	-0,0955	0,1165
31	-0,3628	0,1509	0,2274	-0,0854	-0,1511	0,3785	0,0331	-0,0126	0,2605	0,5747	-0,0049	0,1728	-0,1485	0,2369	0,2274	-0,1514	0,1752
32	-0,0428	0,0988	-0,0930	-0,1624	0,0985	-0,0618	0,0343	0,0765	0,1636	0,8771	0,0819	0,0169	-0,0406	-0,0498	-0,0503	0,2849	0,1404
33	0,5934	0,2309	-0,0885	0,2306	0,1915	-0,3786	0,1170	0,3712	-0,0794	0,0733	-0,1353	0,0070	-0,1923	-0,0077	-0,0257	0,0934	-0,0330
34	0,9148	-0,0781	-0,0099	0,0015	0,1784	-0,0910	0,0177	0,0386	0,1033	0,0909	0,2806	0,4207	-0,0932	0,1765	-0,0229	0,0181	-0,1351
35	0,0343	0,0111	-0,0099	0,0461	-0,0329	0,1748	-0,0025	0,7778	0,1014	0,0492	0,0289	0,1423	-0,0785	-0,0689	-0,0051	0,0423	-0,8864
36	0,1607	0,0782	0,0051	-0,2053	0,2330	-0,0245	-0,0759	-0,0043	0,1014	0,0492	0,0289	0,1423	-0,0785	-0,0689	-0,0051	0,0423	-0,8864
37	-0,0488	0,2252	0,2008	-0,1881	0,2747	0,0347	-0,0901	0,6029	0,3500	0,1760	-0,0609	-0,0582	0,1058	0,2962	0,0584	0,0971	0,3080
38	0,1369	0,0152	-0,2606	0,0094	0,2078	-0,0853	0,2550	0,4805	0,5136	-0,0761	0,5316	-0,0485	0,0240	-0,1173	0,0046	0,1280	0,2615
39	0,0325	-0,4788	-0,3726	0,1521	-0,0897	-0,4567	-0,0222	0,1030	0,1680	0,0196	0,5472	-0,0043	0,0048	-0,0521	0,0006	-0,0092	0,1090
40	0,1103	0,1216	0,1321	-0,1520	0,1239	0,0904	0,0949	-0,1043	-0,2336	0,0482	-0,0231	-0,1187	0,8499	-0,1710	0,2062	0,1410	0,0261
41	-0,0571	0,3389	-0,0133	0,0459	-0,0498	0,7827	0,0637	0,3299	0,0951	0,0395	-0,0650	0,3091	0,1061	-0,0276	-0,0784	-0,0695	0,0824
42	0,6502	-0,3633	0,2812	-0,1161	-0,0675	0,0723	-0,3458	-0,1378	0,1052	0,2186	-0,0372	0,1313	0,2835	-0,0673	-0,0667	0,1311	-0,0661

Matriz de Carregamentos Fatoriais Rotacionados

Continuação

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
43	0,4394	0,1183	0,5003	0,0866	0,0325	0,0270	0,0062	-0,0058	0,1836	0,1831	0,2629	-0,1101	0,1724	0,3190	0,2707	0,3504	0,1016
44	0,3905	0,2673	0,1683	-0,0405	0,6633	-0,0298	-0,0815	0,2558	0,2695	0,2075	-0,0451	0,1165	0,1040	-0,0533	0,2826	-0,0250	-0,1836
45	0,8770	-0,0172	0,1208	-0,2882	0,0019	-0,1122	0,1855	0,0315	0,0221	-0,1818	-0,0874	0,1685	0,0738	0,0837	-0,0313	-0,0215	0,0112
46	0,1684	-0,3664	0,0812	-0,0190	0,0281	-0,1483	0,7904	0,0297	0,0764	-0,0602	0,2505	-0,0055	-0,1349	-0,1818	-0,0292	-0,0736	0,1621
47	0,5020	0,3700	-0,0017	0,4276	-0,0792	0,0890	0,0663	0,5054	-0,1509	0,1898	0,1271	0,1149	-0,0357	0,2344	-0,0002	-0,0796	-0,0796
48	0,1504	0,0218	0,0088	0,1117	0,0765	-0,1774	0,0750	-0,1427	-0,1104	0,0613	-0,1304	-0,1032	0,1136	0,0028	0,0497	0,8642	-0,0840
49	0,3377	0,0016	0,0031	0,1736	0,2527	-0,0523	0,2741	0,2517	0,2749	0,2131	0,2792	0,1629	0,1364	-0,1079	-0,0961	0,6249	0,0549
50	-0,0786	-0,0896	0,2054	0,3416	-0,4259	-0,0156	-0,0456	0,1591	-0,0063	0,5248	0,4404	0,1678	0,1160	0,0548	0,0927	-0,1535	-0,2174
51	-0,1414	0,6859	-0,1967	0,0859	-0,3356	0,1752	-0,4022	0,0051	-0,1577	0,2046	-0,0158	0,1771	0,1795	0,0809	0,0729	0,1044	0,0180
52	-0,2575	-0,0464	-0,0200	-0,1175	-0,0297	-0,0196	-0,1332	-0,1949	-0,0707	0,1302	0,0232	0,0870	0,0977	-0,8838	0,0163	0,0220	-0,0770
53	-0,1872	0,2176	0,1749	0,3808	-0,5414	0,1744	-0,4178	0,1405	-0,1756	-0,0901	0,0386	-0,1569	-0,3535	-0,0192	0,1332	-0,0506	-0,0152
54	-0,2297	-0,8093	-0,0409	0,1485	-0,1381	-0,1552	0,1868	-0,0028	-0,1018	-0,2461	0,0791	-0,0950	-0,0665	-0,0371	0,1403	-0,1470	0,1824
55	-0,2536	0,0959	0,0690	0,1923	-0,8243	0,1359	-0,1917	-0,1458	-0,0739	0,0587	-0,1286	0,0225	0,0847	0,0228	0,1517	-0,0015	0,1873
56	-0,1246	0,1113	0,1264	0,5398	-0,0430	-0,0791	0,4094	0,3856	0,4665	-0,0141	-0,2120	-0,1206	0,1771	-0,0407	0,0151	-0,1424	0,0424
57	0,4099	-0,2944	-0,0170	-0,4149	0,0150	0,1510	-0,3342	-0,1931	0,0150	0,5222	-0,0281	0,2153	0,0725	0,0581	0,2171	-0,0540	-0,0644
58	0,5671	0,3548	-0,0209	0,1863	-0,0868	0,1906	0,2678	0,2606	0,2728	0,2340	-0,3069	-0,0458	-0,0447	0,1529	0,0037	0,0627	0,2062
59	0,1490	-0,3866	0,8852	0,0388	0,0534	-0,0935	-0,0773	-0,0935	-0,0640	-0,0765	0,0065	0,1555	0,0198	-0,0563	0,0404	-0,0574	0,0434
60	0,0893	-0,1086	0,8835	0,1518	-0,0773	0,0884	0,1146	0,1651	0,2346	-0,0304	-0,2181	-0,0730	-0,0425	-0,0066	0,0633	0,0112	0,0201
61	0,6810	0,0891	0,0567	-0,0264	0,5848	-0,2749	0,0497	0,1262	-0,0419	0,1140	0,0323	-0,1116	-0,0830	0,0864	-0,0604	0,1916	-0,0046
62	-0,6602	0,1027	0,2106	0,0481	-0,4598	0,0647	0,0669	-0,1031	-0,0951	0,2039	-0,1351	0,0980	-0,0547	-0,2504	0,1360	-0,2383	0,0875
63	0,4783	0,1787	0,4899	0,3160	-0,0486	-0,0817	0,1158	0,2356	0,1070	0,1815	-0,0348	0,1474	0,4793	0,0614	0,0909	0,0279	0,0075
64	0,1159	0,0200	-0,0883	0,2601	-0,2110	0,1625	-0,0126	-0,0453	-0,3395	0,2152	-0,3302	0,0431	0,3974	0,1378	0,5922	-0,1099	-0,1113
65	0,4020	0,0887	0,1543	0,2462	-0,1186	-0,1997	-0,4358	0,2066	-0,2888	0,0459	-0,3458	-0,2163	0,2142	0,1180	0,1332	-0,0377	-0,3032
66	-0,0624	0,1433	0,0033	0,0606	-0,1123	0,0121	0,0770	-0,0380	-0,9134	-0,0078	-0,0626	-0,0285	0,2104	-0,0595	0,1518	0,0460	0,1643
67	-0,3038	0,0634	-0,4033	0,0673	0,0483	-0,0964	0,2001	-0,5371	0,0725	-0,3272	-0,1528	-0,2194	0,2924	-0,0189	0,0793	0,0458	0,2153
68	-0,0466	0,3423	0,0875	-0,3184	0,2671	0,1354	-0,0134	-0,0247	-0,0746	-0,5186	-0,3198	-0,1831	-0,2959	0,3197	-0,3465	0,4260	0,0809
69	-0,2867	0,2450	-0,1792	0,2038	0,1307	0,4304	-0,1232	-0,0746	-0,5186	0,1803	-0,3381	-0,2238	0,1336	0,0256	0,0500	-0,1902	0,1332
70	-0,1677	0,4410	-0,2149	0,4050	-0,0988	0,5391	-0,0698	0,1473	0,1319	0,0996	-0,0929	-0,1692	-0,1831	-0,0913	-0,1774	0,0064	0,1141
71	-0,0365	-0,0247	0,8398	0,1232	-0,1716	-0,0929	0,1473	0,0186	-0,5186	0,0320	0,1952	0,1205	-0,0725	0,1050	0,2337	0,0896	0,2162
72	-0,1003	-0,2135	0,1726	0,5008	0,4184	-0,3691	-0,1465	0,3777	0,1281	0,0681	0,1952	0,1564	-0,1087	-0,0264	0,0430	0,0560	-0,2699
73	0,0978	0,2897	-0,2880	-0,3409	0,1212	0,2628	0,2663	-0,1304	0,4680	-0,0638	-0,0820	0,0558	-0,2327	0,1408	0,0454	-0,0088	0,4438
74	-0,2307	0,8967	-0,1331	0,0866	-0,0880	0,1029	0,0644	-0,0864	0,0446	-0,0240	-0,1516	-0,0432	-0,0139	0,0293	0,0182	-0,0908	0,2216
75	-0,1826	0,2965	-0,1819	-0,1652	0,3787	0,0074	0,0889	-0,0327	0,1753	0,6225	0,1880	-0,0463	-0,1508	0,1805	0,0182	-0,0457	-0,2407
76	0,1592	-0,0159	0,2281	-0,1652	-0,2105	0,0014	-0,0200	0,1626	-0,3740	0,1830	0,0957	-0,1113	0,0428	-0,0454	0,1200	0,0283	-0,0942
77	0,0985	-0,0393	0,0035	-0,9122	-0,0130	-0,0414	-0,0031	-0,0206	0,1546	0,1830	0,0957	-0,1113	0,0428	-0,0454	0,1200	0,0283	-0,0942
78	-0,0991	0,4162	-0,0979	-0,6966	0,0172	-0,0716	0,1398	0,0147	-0,1414	0,2540	0,0223	-0,0503	0,1695	-0,1034	0,0816	-0,2658	-0,2575
79	-0,1041	0,2764	0,0472	0,1312	-0,3553	0,2315	-0,1918	-0,2886	-0,1355	-0,0372	-0,0019	-0,6654	-0,2006	0,3073	-0,0440	-0,0077	0,0520
80	0,1626	0,1197	-0,1301	0,1355	0,0368	0,2314	-0,0467	-0,0215	-0,3526	0,7149	0,0899	0,2136	0,1644	-0,1286	0,1284	-0,0485	-0,3224

ANEXO 07 – CÓDIGOS

a) Método de Householder

```

' Calcula Matriz Tridiagonal

For i = n To 2 Step -1
  l = i - 1
  h = 0
  mScale = 0
  If l > 1 Then
    For k = 1 To l
      mScale = mScale + Abs(mCorrelacao(i, k))
    Next k
    If mScale = 0 Then
      e(i) = mCorrelacao(i, l)
    Else
      For k = 1 To l
        mCorrelacao(i, k) = mCorrelacao(i, k) / mScale
        h = h + mCorrelacao(i, k) ^ 2
      Next k
      f = mCorrelacao(i, l)
      g = -Abs(Sqr(h)) * Sgn(f)
      e(i) = mScale * g
      h = h - f * g
      mCorrelacao(i, l) = f - g
      f = 0
      For j = 1 To l
        mCorrelacao(j, i) = mCorrelacao(i, j) / h
        g = 0
        For k = 1 To j
          g = g + mCorrelacao(j, k) * mCorrelacao(i, k)
        Next k
        For k = j + 1 To l
          g = g + mCorrelacao(k, j) * mCorrelacao(i, k)
        Next k
        e(j) = g / h
        f = f + e(j) * mCorrelacao(i, j)
      Next j
      hh = f / (h + h)
      For j = 1 To l
        f = mCorrelacao(i, j)
        g = e(j) - hh * f
        e(j) = g
        For k = 1 To j
          mCorrelacao(j, k) = mCorrelacao(j, k) - f * e(k) - g *
mCorrelacao(i, k)
        Next k
      Next j
    End If
  Else
    e(i) = mCorrelacao(i, l)
  End If
  d(i) = h
Next i
d(1) = 0
e(1) = 0
For i = 1 To n
  l = i - 1
  If d(i) <> 0 Then

```

```

For j = 1 To l
  g = 0
  For k = 1 To l
    g = g + mCorrelacao(i, k) * mCorrelacao(k, j)
  Next k
  For k = 1 To l
    mCorrelacao(k, j) = mCorrelacao(k, j) - g * mCorrelacao(k, i)
  Next k
Next j
End If
d(i) = mCorrelacao(i, i)
mCorrelacao(i, i) = 1
For j = 1 To l
  mCorrelacao(i, j) = 0
  mCorrelacao(j, i) = 0
Next j
Next i

```

b) Método QL

' Calcula autovalores e autovetores da matriz tridiagonal

```

For i = 2 To n
  e(i - 1) = e(i)
Next i
e(n) = 0
For l = 1 To n
  iter = 0
  Do
    DONE = True
    For M = l To n - 1
      DD = Abs(d(M)) + Abs(d(M + 1))
      If Abs(e(M)) + DD = DD Then Exit For
    Next M
    If Abs(e(M)) + DD <> DD Then M = n
    If M <> l Then
      If iter = 30 Then
        MsgBox "Muitas iterações."
        Exit Sub
      End If
      iter = iter + 1
      g = (d(l + 1) - d(l)) / (2 * e(l))
      R = Sqr(g ^ 2 + 1)
      g = d(M) - d(l) + e(l) / (g + Abs(R) * Sgn(g))
      s = 1
      c = 1
      p = 0
      For i = M - 1 To l Step -1
        f = s * e(i)
        b = c * e(i)
        If Abs(f) >= Abs(g) Then
          c = g / f
          R = Sqr(c ^ 2 + 1)
          e(i + 1) = f * R
          s = 1 / R
          c = c * s

```

```

Else
  s = f / g
  R = Sqr(s ^ 2 + 1)
  e(i + 1) = g * R
  c = 1 / R
  s = s * c
End If
g = d(i + 1) - p
R = (d(i) - g) * s + 2 * c * b
p = s * R
d(i + 1) = g + p
g = c * R - b
For k = 1 To n
  f = Z(k, i + 1)
  Z(k, i + 1) = s * Z(k, i) + c * f
  Z(k, i) = c * Z(k, i) - s * f
Next k
Next i
d(1) = d(1) - p
e(1) = g
e(M) = 0
DONE = False
End If
Loop While Not DONE
Next l

```

c) Inversão de Matrizes

```

' Para encontrar a inversa de mTempInv
' Programa para inversão de matrizes
' pelo método da eliminação

ReDim ms(NumFat, 2 * NumFat)
ReDim mm(NumFat, 2 * NumFat)
ReDim mTemp(NumFat, nVar)
ReDim mTemp2(NumFat, nVar)
ReDim vt(NumFat)

For X = 1 To NumFat
  For Y = 1 To NumFat
    mm(X, Y) = mTempInv(X, Y)
  Next Y
Next X

' S0
For j = 1 To NumFat
  ' S1
  For i = 1 To NumFat
    ms(j, i) = mm(j, i)
  Next i
  ' S2
  For i = NumFat + 1 To 2 * NumFat
    If i - NumFat <> j Then
      mm(j, i) = 0
    Else
      mm(j, i) = 1
    End If
  Next i
Next j

```

```

        End If
    Next i
    ' S3
Next j
' S4
delta = 1
For k = 1 To NumFat
    ' S5
    p = mm(k, k)
    For i = 1 To NumFat
        vt(i) = mm(i, k)
    Next i
    ' S6
    For j = k To 2 * NumFat
        ' S7
        mm(k, j) = mm(k, j) / p
        For i = 1 To NumFat
            ' S8
            If k <> i Then
                mm(i, j) = mm(i, j) - vt(i) * mm(k, j)
            End If
        Next i
        ' S9
    Next j
    ' S10
    If k < NumFat Then
        delta = delta * mm(k + 1, k + 1)
    End If
Next k
' S11
For i = 1 To NumFat
    ' S12
    For j = 1 To NumFat
        ' S13
        mm(i, j) = 0
        ' S14
        For k = 1 To NumFat
            mm(i, j) = mm(i, j) + ms(i, k) * mm(k, j + NumFat)
        Next k
        ' S15
    Next j
    ' S16
Next i
' S17

```

d) Rotação Varimax

`Matrizes:

L(número de variáveis, número de fatores)

Lrot (número de variáveis, número de fatores)

C(número de variáveis, número de fatores)

T(número de fatores, número de fatores)

D(número de fatores, número de fatores)

B(número de fatores, número de fatores)

A(número de fatores, número de fatores)

Antes de começar a rotação vamos normalizar a matriz nas linhas

```

Ln/sqr(somatório(i(n,1-NumCol)^2)

ReDim mVarimax_L(nVar, NumFat)
ReDim mNormal_L(nVar)
ReDim mVarimax_Lrot(nVar, NumFat)
ReDim mVarimax_C(nVar, NumFat)
ReDim mVarimax_T(NumFat, NumFat)
ReDim mVarimax_D(NumFat, NumFat)
ReDim mVarimax_B(NumFat, NumFat)
ReDim mVarimax_A(NumFat, NumFat)

`Matrizes para cálculos intermediários
ReDim mVarimax_Temp1(nVar, NumFat)
ReDim mVarimax_Temp2(NumFat, NumFat)

`Usando o resultado de mm anterior
For X = 1 To NumFat
mNormal_L(X) = 0
  For Y = 1 To NumFat
    Z(X, Y) = 0
    If X = Y Then
      Z(X, Y) = 1
    End If
    mVarimax_L(X, Y) = mm(X, Y)
    mVarimax_Lrot(X, Y) = mm(X, Y)
    mNormal_L(X) = mNormal_L(X) + mm(X, Y) ^ 2
  Next Y
Next X

Início das iterações

Do While True
  Fazendo c=1^3
  For X = 1 To nVar
    For Y = 1 To NumFat
      mVarimax_C(X, Y) = mVarimax_Lrot(X, Y) ^ 3
    Next Y
  Next X

  `Encontrando d=somatório(i,1,p,1^2)
  For X = 1 To NumFat
    For Y = 1 To NumFat
      mVarimax_D(X, Y) = 0
      If X = Y Then
        For T = 1 To nVar
          mVarimax_D(X, Y) = mVarimax_D(X, Y)+mVarimax_Lrot(T, Y)^2
        Next T
      End If
    Next Y
  Next X

  `Encontrando B=L'[C-(1/nVar)Lrot.D]

  `Encontrando Lrot.D e armazenando em mVarimax_Temp1
  For Y = 1 To nVar
    For T = 1 To NumFat
      mVarimax_Temp1(Y, T) = 0
      For X = 1 To NumFat

```

```

                mVarimax_Temp1(Y, T) = mVarimax_Temp1(Y, T) +
mVarimax_Lrot(Y, X) * mVarimax_D(X, T)
            Next X
        Next T
    Next Y

    `Encontrando C-(1/nVar)*mVarimax_Temp1 e armazenando
    em mVarimax_Temp1
    For X = 1 To nVar
        For Y = 1 To NumFat
            mVarimax_Temp1(X, Y) = mVarimax_C(X, Y) - (1 / nVar) *
mVarimax_Temp1(X, Y)
        Next Y
    Next X

    `Encontrando B = L' * mVarimax_Temp1
    For Y = 1 To NumFat
        For T = 1 To NumFat
            mVarimax_B(Y, T) = 0
            For X = 1 To nVar
                mVarimax_B(Y, T) = mVarimax_B(Y, T) + mVarimax_L(X, Y) *
mVarimax_Temp1(X, T)
            Next X
        Next T
    Next Y

    `Encontrando A^2 e armazenando em mVarimax_Temp2
    For Y = 1 To NumFat
        For T = 1 To NumFat
            mVarimax_Temp2(Y, T) = 0
            For X = 1 To NumFat
                mVarimax_Temp2(Y, T) = mVarimax_Temp2(Y, T) + mVarimax_B(X,
Y) * mVarimax_B(X, T)
            Next X
        Next T
    Next Y

    `Encontrando autovalores e autovetores de mVarimax_Temp2

    `Calcula Matriz Tridiagonal
    For i = NumFat To 2 Step -1
        l = i - 1
        h = 0
        mScale = 0
        If l > 1 Then
            For k = 1 To l
                mScale = mScale + Abs(mVarimax_Temp2(i, k))
            Next k
            If mScale = 0 Then
                e(i) = mVarimax_Temp2(i, 1)
            Else
                For k = 1 To l
                    mVarimax_Temp2(i, k) = mVarimax_Temp2(i, k) / mScale
                    h = h + mVarimax_Temp2(i, k) ^ 2
                Next k
                f = mVarimax_Temp2(i, 1)
                g = -Abs(Sqr(h)) * Sgn(f)
                e(i) = mScale * g
                h = h - f * g
            End If
        End If
    Next i

```



```

mVarimax_Temp2(i, 1) = f - g
f = 0
For j = 1 To 1
  mVarimax_Temp2(j, i) = mVarimax_Temp2(i, j) / h
  g = 0
  For k = 1 To j
    g = g + mVarimax_Temp2(j, k) * mVarimax_Temp2(i, k)
  Next k
  For k = j + 1 To 1
    g = g + mVarimax_Temp2(k, j) * mVarimax_Temp2(i, k)
  Next k
  e(j) = g / h
  f = f + e(j) * mVarimax_Temp2(i, j)
Next j
hh = f / (h + h)
For j = 1 To 1
  f = mVarimax_Temp2(i, j)
  g = e(j) - hh * f
  e(j) = g
  For k = 1 To j
    mVarimax_Temp2(j, k) = mVarimax_Temp2(j, k) - f * e(k) - g
* mVarimax_Temp2(i, k)
  Next k
Next j
End If
Else
  e(i) = mVarimax_Temp2(i, 1)
End If
d(i) = h
Next i
d(1) = 0
e(1) = 0
For i = 1 To NumFat
  l = i - 1
  If d(i) <> 0 Then
    For j = 1 To 1
      g = 0
      For k = 1 To 1
        g = g + mVarimax_Temp2(i, k) * mVarimax_Temp2(k, j)
      Next k
      For k = 1 To 1
        mVarimax_Temp2(k, j) = mVarimax_Temp2(k, j) - g *
mVarimax_Temp2(k, i)
      Next k
    Next j
  End If
  d(i) = mVarimax_Temp2(i, i)
  mVarimax_Temp2(i, i) = 1
  For j = 1 To 1
    mVarimax_Temp2(i, j) = 0
    mVarimax_Temp2(j, i) = 0
  Next j
Next i

` Inicializa Z para a saída do tred2
For Y = 1 To NumFat
  For X = 1 To NumFat
    Z(Y, X) = mVarimax_Temp2(Y, X)
  Next X

```

```

Next Y

' Calcula autovalores e autovetores da matriz tridiagonal
For i = 2 To NumFat
    e(i - 1) = e(i)
Next i
e(NumFat) = 0
For l = 1 To NumFat
    iter = 0
    Do
        DONE = True
        For M = 1 To NumFat - 1
            DD = Abs(d(M)) + Abs(d(M + 1))
            If Abs(e(M)) + DD = DD Then Exit For
        Next M
        If Abs(e(M)) + DD <> DD Then M = NumFat
        If M <> 1 Then
            If iter = 30 Then
                MsgBox "Muitas iterações." ' mensagem de erro'.
                Exit Sub
            End If
            iter = iter + 1
            g = (d(l + 1) - d(l)) / (2 * e(l))
            R = Sqr(g ^ 2 + 1)
            g = d(M) - d(l) + e(l) / (g + Abs(R) * Sgn(g))
            s = 1
            c = 1
            p = 0
            For i = M - 1 To 1 Step -1
                f = s * e(i)
                b = c * e(i)
                If Abs(f) >= Abs(g) Then
                    c = g / f
                    R = Sqr(c ^ 2 + 1)
                    e(i + 1) = f * R
                    s = 1 / R
                    c = c * s
                Else
                    s = f / g
                    R = Sqr(s ^ 2 + 1)
                    e(i + 1) = g * R
                    c = 1 / R
                    s = s * c
                End If
                g = d(i + 1) - p
                R = (d(i) - g) * s + 2 * c * b
                p = s * R
                d(i + 1) = g + p
                g = c * R - b
                For k = 1 To NumFat
                    f = Z(k, i + 1)
                    Z(k, i + 1) = s * Z(k, i) + c * f
                    Z(k, i) = c * Z(k, i) - s * f
                Next k
            Next i
            d(l) = d(l) - p
            e(l) = g
            e(M) = 0
            DONE = False
        Loop While Not DONE
    Loop While Not DONE
Next l

```

```

        End If
    Loop While Not DONE
Next l

' Cálculo de  $(A1^2)^{-1}$ 
' usando Z = autovetores e d = autovalores
' armazenando resultados temporários em
' mVarimax_Temp2

For Y = 1 To NumFat
    For T = 1 To NumFat
        mVarimax_Temp2(Y, T) = 0
        For X = 1 To NumFat
            If X = T Then
                mVarimax_Temp2(Y, T) = mVarimax_Temp2(Y, T) + Z(Y, X) *
(1 / qr(d(X)))
            End If
        Next X
    Next T
Next Y

For Y = 1 To NumFat
    For T = 1 To NumFat
        mVarimax_A(Y, T) = 0
        For X = 1 To NumFat
            If X = T Then
                mVarimax_A(Y, T) = mVarimax_A(Y, T) + mVarimax_Temp2(Y,
X) * Z(T, X)
            End If
        Next X
    Next T
Next Y

' Atualizando T para a próxima iteração
For Y = 1 To NumFat
    For T = 1 To NumFat
        mVarimax_T(Y, T) = 0
        For X = 1 To NumFat
            mVarimax_T(Y, T) = mVarimax_T(Y, T) + mVarimax_B(Y, X) *
Varimax_A(X, T)
        Next X
    Next T
Next Y

' Atualizando L para a próxima iteração
For Y = 1 To nVar
    For T = 1 To NumFat
        mVarimax_Temp1(Y, T) = 0
        For X = 1 To NumFat
            mVarimax_Temp1(Y, T) = mVarimax_Temp1(Y, T) +
mVarimax_Lrot(Y, X) * mVarimax_T(X, T)
        Next X
    Next T
Next Y

For Y = 1 To nVar
    For T = 1 To NumFat
        mVarimax_L(Y, T) = mVarimax_Lrot(Y, T)
        mVarimax_Lrot(Y, T) = mVarimax_Temp1(Y, T)
    Next T
Next Y

```

Loop

```
' Encontrando o vetor de variâncias específicas
ReDim mVarEsp(n)
For X = 1 To n
    mVarEsp(X) = 1 - (mL(X, 1) ^ 2 + mL(X, 2) ^ 2)
Next X

' Encontrando o vetor do valor médio amostral
ReDim mValMed(nVar)
For X = 1 To nVar
    mValMed(X) = 0
    For Y = 1 To nDad
        mValMed(X) = mValMed(X) + mDados(X, Y)
    Next Y
    mValMed(X) = mValMed(X) / nDad
Next X
```

ANEXO 08 – FUNÇÃO DE AJUDA DO PROGRAMA

FUNÇÃO DE AJUDA DO PROGRAMA

1. Acionando a opção *Ajuda*, tem-se acesso aos principais esclarecimentos sobre o programa de ranqueamento. A entrada de dados é sempre efetuada através de uma matriz. Para tanto, considera-se as variáveis a serem analisadas nas colunas e as observações (Áreas Especialmente Protegidas) nas linhas.
2. O número de variáveis é indeterminado. Está limitado apenas à capacidade da máquina que se utiliza. Para se obter o ranqueamento, as variáveis analisadas devem ser as mesmas para cada observação. Sem isso não é possível aplicar a Análise Fatorial.
3. O usuário pode, caso seja de seu interesse, observar as matrizes intermediárias dos cálculos que fazem parte do desenvolvimento, até se alcançar os escores finais. As matrizes do desenvolvimento são: a matriz de correlação das variáveis, a matriz dos autovalores e autovetores da matriz de correlação, a matriz das communalidades e variâncias específicas das variáveis, a matriz de carregamentos fatoriais, a matriz de carregamentos fatoriais rotacionados e a matriz dos escores fatoriais.
4. Caso o interesse na Análise Fatorial não seja o ranqueamento, mas sim, a interpretação dos fatores e o agrupamento das variáveis em cada fator, pode-se obter essa interpretação com a matriz de carregamentos fatoriais rotacionados. A matriz de carregamentos fatoriais rotacionados é a matriz de carregamentos fatoriais após uma rotação Varimax.
5. Os escores brutos e os escores padronizados podem ser armazenados na memória acionando a opção *Armazenar Escores*. Caso os escores de anos anteriores sejam requeridos, podem ser obtidos acionando *Escores de Anos Anteriores*.
6. O programa apresenta ainda uma *Tábua de Avaliação*, na qual podem ser colocadas a identificação da pessoa que realizou a coleta de dados, a localização da unidade de conservação, a qualificação da Área Especialmente Protegida e a responsabilidade técnica do avaliador. Todas essas informações também podem ser gravadas na memória.
7. Acionando *Configuração*, tem-se acesso, além da Unidade de Conservação e descrição das variáveis, ao intervalo ou faixa em se deseja situar os escores padronizados. Desta forma, qualquer que seja o tipo de Área Especialmente Protegida, pode-se padronizar os escores da forma mais conveniente possível, de acordo com o interesse do usuário.
8. Para maiores esclarecimentos, basta acessar a opção *Ajuda*.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 AFONSO, C.; FURTADO, E. M.; COSTA, M. B. F. **Aplicações de Análise Multivariada**. Curitiba: UFPR, 1999. 83 p. (lista de exercícios).
- 2 ANDERSON, T. W. **An introduction to multivariate statistical analysis**, New York: John Wiley & Sons INC., 1958. 374 p.
- 3 ATKINSON, Cyril P. **Programação e métodos computacionais**, Rio de Janeiro: LTC 1977. 665 p.
- 4 BARTLETT, M. S. A note on multiplying factors for various chi squared approximations. **Journal of the Royal Statistical Society (B)**, 16, 1954. 296-298 p.
- 5 BEZERRA, M. J.; “JOTA”, J. C. P. **Matemática**. São Paulo: Scipione, 1994. 583p.
- 6 CHANG, M. Y. **Faxinais no Paraná**. Londrina: IAPAR, 1988. 20 p. (Informe de Pesquisa, n. 80).
- 7 _____. **Sistema Faxinal. Uma forma de organização camponesa em desagregação no Centro-Sul do Paraná**, Londrina: IAPAR, 1988, 123 p. (Boletim Técnico, n. 22).
- 8 CHAVES NETO. **Análise multivariada aplicada à pesquisa**. Curitiba: UFPR, 1998, 74 p. (notas de aula).
- 9 CIGOLINI, Adilar; MELLO Laércio de; LOPES, Nelci; PARANÁ – **Quadro natural, transformações territoriais e economia**. Curitiba: Renascer, 1998. 128 p.
- 10 COOLEY, William W.; LOHNES, Raul R. **Multivariate data analysis**. New York: John Wiley & Sons INC. 1971. 364 p.
- 11 DOMINGUES, Zilna H. **Hierarquização de faxinais inscritos no cadastro estadual de unidades de conservação e uso especial, visando o ICMS ecológico**. Curitiba, 1999, Dissertação (Mestrado) – Setor de Ciências Agrárias, Curso de Engenharia Florestal, Universidade Federal do Paraná, 143 p.
- 12 EDWARDS Jr.; C. H. PENNEY, David E. **Introdução à álgebra linear**. Rio de Janeiro: Prentice Hall do Brasil, 1998, 406 p.
- 13 GOLUB, G.; VAN LOAN, C. **Matrix computations**. The Johns Hopkins University Press.
- 14 HALVORSON, Michael. **Microsoft visual basic 5 – step by step**. São Paulo: Makron Books, 1997. 379 p.
- 15 JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4. ed. New Jersey: Prentice Hall, Englewood cliffs, 1998, 816 p.

- 16 KAYSER, Henry F. **The varimax criterion for analytic rotation in factor analysis.** Psychometrika, v. 23, n. 3, University of Illinois, 1958.
- 17 KUO, Shan S. **Numerical methods and computers**, London: Addison-Wesley Publishing Company, 1957, 381p.
- 18 LAWLEY, D. N.; MAXWELL, A. E. **Factor analysis as a statistical method.** New York: American Elsevier Publishing Company, 1971, 153p.
- 19 LOIOLA, Caetano R. A. **Rotinas matemáticas em BASIC para micros.** Rio de Janeiro: Editora Campus, 1985. 232 p.
- 20 LOUREIRO, Wilson. **ICMS Ecológico em perguntas e respostas.** Governo do estado do Paraná, 1998, 34 p.
- 21 MacLANE S.; BIRKHOFF G. **Algebra.** New York: The Macmillan Company, 1967, 598 p.
- 22 MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis.** Academic press INC. New York, 1979. 279 p.
- 23 McCALLA, Thomas Richard, **Introduction to numerical methods and FORTRAN programming**, New York: John Wiley & Sons Inc., 1967. 359p.
- 24 NASH, J. C. **Compact numerical methods for computers: linear algebra and function minimisation.** 2. ed. New York: Adam Hilger, 1990. 278 p.
- 25 PRESS, S. James. **Applied Multivariate analysis.** Chicago: Holt, Rinehart and Winston, INC. 1971. 520 p.
- 26 RENCHER, Alvin C. **Methods of multivariate analysis.** New York: Wiley InterScience, 1995. 627 p.
- 27 _____. **Multivariate statistical inference and applications**, New York: Wiley InterScience, 1998. 557 p.
- 28 STEINBRUCH, Alfredo; WINTERLE, Paulo. **Álgebra linear.** São Paulo: McGraw-Hill, 1987. 583 p.
- 29 TATSUOKA, Maurice M. **Multivariate data analysis.** New York: John Wiley & Sons, INC, 1971. 310 p.
- 30 VETTERLING, William T. et al. **Numerical recipes example book (FORTRAN).** 2. ed. London: Cambridge University Press, 1992. 245 p.