

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA E
MODELAGEM QUANTITATIVA**

**ANÁLISE MULTIVARIADA DE DADOS
SOCIOECONÔMICOS: UM RETRATO DA
MODERNIZAÇÃO AGRÍCOLA NO RIO GRANDE DO
SUL**

MONOGRAFIA DE ESPECIALIZAÇÃO

Rafael Pentiado Poerschke

**Santa Maria, RS, Brasil
2007**

**ANÁLISE MULTIVARIADA DE DADOS
SOCIOECONÔMICOS:**

**UM RETRATO DA MODERNIZAÇÃO AGRÍCOLA NO RIO GRANDE
DO SUL**

por

Rafael Pentiado Poerschke

Monografia apresentada ao Curso de Especialização do Programa de Pós-Graduação em Estatística e Modelagem Quantitativa, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Especialista em Estatística e Modelagem Quantitativa,**

Orientador: Prof. Ms. Fernando de Jesus Moreira Júnior

Santa Maria, RS, Brasil

2007

© 2007

Todos os direitos autorais reservados a Rafael Pentiado Poerschke. A reprodução de partes ou do todo deste trabalho só poderá ser com autorização por escrito do autor.

Endereço: Rua Venâncio Aires, n.920/502, Bairro Centro, Santa Maria, RS, 97010-000

Fone (0xx)55 32221077; End. Eletr: faecohaiti@yahoo.com.br

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Departamento de Estatística
Curso de Especialização em Estatística e Modelagem Quantitativa**

A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

**ANÁLISE MULTIVARIADA DE DADOS SOCIOECONÔMICOS: UM
RETRATO DA MODERNIZAÇÃO AGRÍCOLA DO RIO GRANDE DO
SUL**

elaborada por
Rafael Pentiado Poerschke

como requisito parcial para a obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa,

COMISSÃO EXAMINADORA:

Fernando de Jesus Moreira Júnior, Prof. Ms. (UFSM)
(Presidente/Orientador)

Rita Inês Pauli Prieb, Prof. Dra. (UFSM)

Ivanor Müller, Prof. Dr. (UFSM)

Santa Maria, 19 de novembro de 2007.

Aos meus pais,
por me guiarem e servirem de exemplo de vida;
ao meu Avô Amaury,
que já não se encontra entre nós;

DEDICO.

AGRADECIMENTOS

Aos meus pais, Maria de Fátima e Paulo R. Cardoso Poerschke, que mais uma vez, com auxílio e compreensão contribuíram para que este trabalho se concretizasse. Agradeço pelo carinho, dedicação e o constante incentivo para continuar estudando. Como não poderia deixar de lembrar, agradeço também, por estarem em todos os momentos que tornaram claro o sentido da palavra família.

Ao meu orientador, Fernando de Jesus Moreira Júnior, pela confiança em mim depositada e por ter realizado seu papel, hora como um grande amigo, hora como um exigente mestre.

À minha amiga e professora Rita Inês Pauli Prieb, por ter aceitado fazer parte da banca desta Monografia, também, pelo incentivo e auxílio.

Ao professor Ivanor Müller, mais uma vez, um grande amigo que fiz durante o curso, bem como, um excelente professor.

Ao professor Adriano Mendonça de Souza que de forma direta ou indireta contribuiu para a realização desta Monografia.

Ao professor Cláilton Ataíde de Freitas pela amizade e a quase total co-orientação nesta Monografia.

Não poderia deixar de agradecer também a “Sandrinha da Secretaria”, pela amizade, atenção e sua constante disponibilidade para ajudar nos problemas burocráticos surgidos durante o curso.

À Carolina Seeger Bortoluzzi pelo amor e companheirismo durante a realização deste trabalho.

A todos os professores presentes em minha formação do curso de Estatística e Modelagem Quantitativa, pelo acolhimento e auxílio no decorrer de minha passagem pelo departamento.

À Universidade Federal de Santa Maria, pela excelência do ensino público.

Enfim, aos meus verdadeiros amigos e pessoas sinceras que me cercam. A vocês que sempre estiveram do meu lado e me agüentaram tanto nos momentos tristes e de extremo mau humor, como também nos melhores momentos da minha vida, o meu muito obrigado.

Não há nada mais grotesco,
[...] do que se levantar cedo.
O ser humano precisa dormir.
(Franz Kafka)

RESUMO

Monografia de Especialização
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria, RS, Brasil

ANÁLISE MULTIVARIADA DE DADOS SOCIOECONÔMICOS: UM RETRATO DA MODERNIZAÇÃO AGRÍCOLA DO RIO GRANDE DO SUL

Autor: Rafael Pentiado Poerschke
Orientador: Prof. Ms. Fernando de Jesus Moreira Júnior
Data e Local da Defesa: Santa Maria, 19 de novembro de 2007.

O processo de modernização da agricultura brasileira desencadeado pelos governo militares teve forte intervenção do Estado através de políticas agrícolas. Como é fato em todo o Brasil, bem como no Rio Grande do Sul (RS), essa evolução tecnológica não se deu de forma uniforme na agropecuária Brasileira. As regiões que exploravam produtos dinâmicos e potencialmente geradores de divisas conseguiram, especialmente, no período de crédito farto e subsidiado, um maior dinamismo econômico. Dessa forma, a utilização de duas ferramentas da estatística multivariada – Análise Fatorial e de Agrupamentos –, permitiu classificar os municípios dos Conselhos Regionais de Desenvolvimento (COREDEs) agrícolas do Rio Grande do Sul quanto ao grau de modernização que se encontravam em 1995/96. Tendo sido utilizados dados censitários *cross-section*, do Censo Agropecuário de 1995/96 do Instituto Brasileiro de Geografia e Estatística (IBGE). No, Rio Grande do Sul, o então estágio de modernização agrícola aponta para uma forte concentração da agricultura moderna e da agropecuária no grupo formado pelos municípios do COREDE Fronteira Oeste. Os demais municípios agrupados não demonstraram padrão local de agrupamento. Contudo, o grupo 2 tem uma situação mais estável e, o grupo 3 demonstrou sinais de transição da agricultura mais tradicional para um padrão diferenciado, caracterizado pela inovação. Revelando, um quarto grupo com pior grau de modernização da agricultura do Estado, sem exploração intensiva da área explorada para pecuária e, em virtude de seu elevado tamanho, carecendo de uma atenção especial para explorar possíveis potencialidades, em vista, de apontar para um padrão de agricultura tradicional.

Palavras-chave: Modernização Agrícola, Análise Multivariada, COREDEs, Rio Grande do Sul.

ABSTRACT

Monografia de Especialização
Programa de Pós-Graduação em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria, RS, Brasil

ANÁLISE MULTIVARIADA DE DADOS SOCIOECONÔMICOS: UM RETRATO DA MODERNIZAÇÃO AGRÍCOLA DO RIO GRANDE DO SUL

(MULTIVARIATE ANALYSIS ON SOCIO-ECONOMIC DATA: A PORTRAIT OF THE AGRICULTURAL MODERNIZATION OF RIO GRANDE DO SUL)

Author: Rafael Pentiado Poerschke

Advisor: Prof. Ms. Fernando de Jesus Moreira Júnior

Date and Place of defense: Santa Maria, November 19 of 2007.

The process of modernization of Brazilian agriculture sparked by military administrations had a strong intervention by the State through farm policies. As seen all over Brazil, as well as in Rio Grande do Sul (RS), this technological evolution did not happen in a uniform way in Brazilian agribusiness. The regions that explored dynamic and potential revenues generating products were able to get, especially in the period of broad and subsidized credit, a higher economic dynamism. So the use of two multivariate statistics tools – Factor Analysis and Cluster Analysis –, made possible the rating of the cities of the agricultural Regional Development Councils (COREDEs) of Rio Grande do Sul about their degree of modernization in the 1995/96 period. Taking into account cross-section census data, from the Agricultural Census of 1995/96 by the Brazilian Institute of Geography and Statistics (IBGE). In Rio Grande do Sul, the stage of agricultural modernization of the analyzed period points to a strong concentration of modern agriculture and animal husbandry in the group made up by the cities of the Western Border COREDE. The other clustered cities did not show a local pattern of clustering. However, group 2 has a steadier situation, and group 3 showed signs of transition from a more traditional agriculture to a differentiated pattern, characterized by innovation. Revealing a fourth group with the worst degree of agriculture modernization of the State, without intensive exploration of the area destined to livestock and, in the wake of its high size, in need of special attention to explore possible potentialities, aiming to point to a pattern of traditional agriculture.

Key words: Agricultural Modernization, Multivariate Analysis, COREDEs, Rio Grande do Sul.

LISTA DE FIGURAS

Figura 1 – Distância mínima entre grupos	54
Figura 2 – Dendrograma referente ao agrupamento pelo método de ligação simples através da distância Euclidiana	58
Figura 3 – Distância máxima entre grupos.....	59
Figura 4 – Dendrograma referente ao agrupamento pelo método de Ward	65
Figura 5 – Representação gráfica dos autovalores e autovetores	75
Figura 6 – Gráfico das distâncias de fusão dos grupos – maior salto.....	96
Figura 7 – Dendrograma – maior salto	97
Figura 8 – Gráfico das distâncias de fusão dos grupos.....	98
Figura 9 – Dendrograma.....	100
Figura 10 – Representação gráfica dos autovalores	106

LISTA DE TABELAS

Tabela 1 – Taxa real de juros do Crédito Rural no Brasil – 1969/86	30
Tabela 2 – Inflação e a evolução dos recursos dos financiamentos concedidos aos produtores e cooperativas no Brasil – 1969/1994.....	33
Tabela 3 – Despesa Total da União, PIB, DGFA, participação do gasto em agricultura nos gastos totais – 1990/2002	41
Tabela 4 – Matriz de dados n indivíduos e p variáveis	51
Tabela 5 – Número de indivíduos e suas respectivas variáveis.....	54
Tabela 6 – Resultado para o teste KMO e a adequabilidade dos dados	68
Tabela 7 – Observações relativas às duas variáveis, levantadas em cinco indivíduos.....	70
Tabela 8 – Estatística descritiva relativa às duas variáveis, levantadas em cinco indivíduos	70
Tabela 9 – Matriz de dados originais e a nova matriz gerada – MS.....	82
Tabela 10 – Resumo da análise de Componentes Principais pela matriz variância-covariância	83
Tabela 11 – Matriz de dados originais e a nova matriz gerada – MR.....	89
Tabela 12 – Resumo da análise de componentes principais pela matriz de correlação	90
Tabela 13 – Orientação para validação das cargas fatoriais com base no tamanho da amostra	93
Tabela 14 – Matriz de correlação entre as variáveis originais normalizadas	101
Tabela 15 – Matriz de correlação ente as variáveis originais normalizadas e as Componentes Principais – significativas $> 0,50$	103
Tabela 16 – Escores fatoriais para os grupos de I a IV	104
Tabela 17 – Raízes características – <i>eigenvalues</i> /autovalores – da matriz de correlação simples (17×17) para os determinados COREDEs – 1995/96.....	107
Tabela 18 – Percentagem da variância explicada por cada fator após a rotação ortogonal ...	108
Tabela 19 – Cargas fatoriais , communalidades e percentagem da variância explicada por cada fator após a rotação, na extração de quatro fatores em 17 variáveis.....	108
Tabela 20 – Escores fatoriais para os grupos de I a IV após rotacionados os fatores	111
Tabela 21 – Classificação em ordem decrescente do nível tecnológico, participação percentual no valor da produção agropecuária, animal e vegetal para os quatro grupos homogêneos dos COREDEs agrícolas do Estado do Rio Grande do Sul	113

LISTA DE QUADROS

Quadro 1 – Variáveis que compõe o estudo	23
Quadro 2 – Processo de agrupamento segundo o método de Ward – Passo 1-2.....	62
Quadro 3 – Processo de agrupamento segundo o método de Ward – Passo 3	63
Quadro 4 – Processo de agrupamento segundo o método de Ward – Passo 4	63
Quadro 5 – Processo de agrupamento segundo o método de Ward – Passo 5	64

LISTA DE ABREVIATURAS E SIGLAS

AA –	Análise de Agrupamento
ACP –	Análise das Componentes Principais
ADUBOS –	Número de estabelecimentos que utilizam adubos químicos/total de estabelecimentos no RS
AF –	Análise Fatorial
AGF –	Aquisição do Governo Federal
ASSPEC –	Assistência Técnica Especializada na Produção Animal/total de estabelecimentos no RS
ASSAGR –	Assistência Técnica Especializada na Produção Vegetal/total de estabelecimentos no RS
AREXPLO –	Área explorada nos municípios/total da área do RS
BB –	Banco do Brasil (BB)
BNCC –	Banco Nacional de Crédito Cooperativo
BNDES –	Banco Nacional Econômico e Social
BASPH –	<i>Barlett of Sphericity</i>
BTN –	Bônus do Tesouro Nacional
CEPEA –	Centro de Estudos Avançados em Economia Aplicada
CIBRAZEM –	Companhia Brasileira de Armazenamento
COBAL –	Companhia Brasileira de Alimentos
CONAB –	Companhia Nacional de Abastecimento
CP –	Componentes Principais
COREDEs –	Conselhos Regionais de Desenvolvimento
DGFA –	Despesas do Governo na Função Agricultura
DISTRG –	Distância do município em relação à cidade portuária de Rio Grande/total das distâncias em relação ao porto
EGF –	Empréstimos do Governo Federal

EMBRATER –	Empresa Brasileira de Extensão Rural
EMBRAPA –	Empresa Brasileira de Pesquisa Agropecuária
FINANC –	Valor dos financiamentos nos municípios/valor total no RS
FEE –	Fundação de Economia e Estatística
GADO –	Efetivo de animais de grande porte do município/total do RS
GPS –	<i>Global Positioning System</i>
IBGE –	Instituto Brasileiro de Geografia e Estatística
IPC –	Índice de Preços ao Consumidor
IRRIGA –	Número de estabelecimentos que utilizam sistemas de irrigação/total de estabelecimentos no RS
KMO –	<i>Kaiser-Meyer-Olkin Measure of Adequacy</i>
MAQCOL –	Números de máquinas para colheita/ total de máquinas para colheita no RS
MP –	Medida Provisória
MR –	Matriz de Correlação
MS –	Matriz de Variância-Covariância
PEA –	População economicamente ativa do município/total da população economicamente ativa no RS
PGPM –	Política de Garantia de Preços Mínimos
PIB –	Produto Interno Bruto
POP –	População rural do município/total da população rural no RS
PROAGRO –	Programa de Garantia da Atividade Agropecuária
RECVEG –	Valor das receitas advindas da venda de produtos vegetais/valor total no RS
RECANI –	Valor das receitas advindas da venda de animais/valor total no RS
RENDA –	Índice bloco renda no município
SNAC –	Sistema Nacional de Crédito Rural
TRATOR –	Número de tratores/total de tratores no RS
URV –	Unidade Real de Valor
VAB –	Valor Adicionado Bruto
VBC –	Valor Básico de Custeio
VALVEG –	Valor da produção vegetal/valor total no RS
VALANI –	Valor da produção animal de grande porte/valor total no RS

LISTA DE ANEXOS

ANEXO A – COREDEs utilizados para o estudo.....	125
ANEXO B – Mancha de Modernização dos municípios dos COREDEs essencialmente agrícolas.....	126

SUMÁRIO

1 INTRODUÇÃO.....	16
1.1 Objetivos.....	19
1.1.1 Objetivo geral.....	19
1.1.2 Objetivos específicos.....	19
1.2 Justificativa.....	20
1.3 Estrutura do trabalho.....	20
2 METODOLOGIA E FONTE DE DADOS.....	21
3 MODERNIZAÇÃO DA AGRICULTURA BRASILEIRA – A ERA DAS MUDANÇAS ESTRUTURAIS.....	24
3.1 Do Desenvolvimento e Modernização da Agricultura à Realidade Inflacionária.....	24
3.1.1 Os Planos Heterodoxos: O comportamento da intervenção do Estado na agricultura - 1986 a 1994.....	26
3.1.2 O Plano Real.....	38
3.2 Mudanças estruturais da agricultura.....	43
3.2.1 A Produção e Produtividade no Rio Grande do Sul.....	44
4 ANÁLISE ESTATÍSTICA MULTIVARIADA DE DADOS.....	49
4.1 Análise de Agrupamento (AA).....	50
4.1.1 As medidas de similaridade ou dissimilaridade.....	50
4.1.2 Os métodos de Análise de Agrupamento.....	53
4.1.2.1 O Método da Ligação Simples.....	53
4.1.2.2 O Método da Ligação Completa.....	58
4.1.2.3 O Método das Médias das Distâncias.....	59
4.1.2.4 O Método do Centróide.....	60
4.1.2.5 O Método de Ward.....	60
4.1.3 Escolha do método e número de grupos.....	65
4.2 Análise fatorial de componentes principais.....	66
4.2.1 Análise de Componentes Principais (ACP).....	67
4.3 Análise Fatorial (AF).....	90
5 ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	94
5.1 Análise de Agrupamento (AA).....	94
5.2 Análise Fatorial (AF).....	101
6 CONCLUSÕES.....	116
REFERÊNCIAS BIBLIOGRÁFICAS.....	118
ANEXO.....	124

1 INTRODUÇÃO

A agricultura brasileira sofreu diversas mudanças estruturais e conjunturais nos últimos cinquenta anos. Da mesma forma, os meios de informação e transporte passaram por igual revolução. Da agricultura tradicional e pastoreio às colheitadeiras guiadas via GPS – *Global Positioning System* –, das caravelas e do precário telégrafo a satélites e navios mercantes que interligam continentes. Mas toda essa panacéia capitalista se deveu e continua devendo à Terceira Revolução Industrial, a partir da invenção do primeiro computador.

Nos países alinhados ao bloco capitalista, a difusão da tecnologia reduziu as distâncias e, gradualmente, aumentou o inter-relacionamento e interdependência. A agricultura, dentro desse desenvolvimento, (re)encontrou sua vocação.

Johnston & Mellor (1961) foram felizes em enumerar as funções que viam para a agricultura em meio à revolução por que passava o mundo capitalista. A agricultura, durante o processo de seu desenvolvimento, seria responsável pela transferência de mão-de-obra para o setor urbano-industrial; continuaria com o fornecimento de alimentos para a população; geração de divisas – exportações do excedente ou de produtos com vantagens comparativas/competitivas – e transferência de poupanças e consumo de bens industriais – complexos agroindustriais, os quais a agricultura estaria integrada. Estes pontos abrangem a essência das funções da agricultura para o desenvolvimento econômico de um País.

O desenvolvimento dos microprocessadores e, em especial, da indústria de *softwares* coadunaram com o aumento das pesquisas em todos os ramos da ciência moderna. Assim como nos meios de transporte, que se tornaram mais eficientes podendo atingir longas distâncias a custos menores. Assim, vantagens comparativas difundidas pelos mais diversos meios de comunicação puderam ser exploradas. Mas, não só com a comunicação e transporte os *softwares* e micro-processadores corroboraram, ainda, a Estatística e Economia se beneficiaram desse salto tecnológico. Programas mais modernos, dispensando a presença de programadores, aliado a computadores com maior poder de processamento possibilitaram o avanço das pesquisas com elevado número de observações, como é o caso de uma análise multivariada de variância.

Doravante, Estados passaram a centralizar informações e sinalizar para a realidade imposta pelo mercado de ajuste fiscal/macroeconômico e credibilidade a ser transmitida ao mercado. Institutos dedicados ao levantamento de dados foram criados ou modernizados

difundindo dados fundamentais e, agora, muito mais confiáveis aos pesquisadores. Censos e amostras de domicílios passaram a fornecer uma quantidade incomensurável de informações.

Tardiamente, o Governo está realizando o Censo Agropecuário de 2007/08 para levantar informações sobre a agricultura e pecuária nacional. São dados essenciais para a economia, contudo, o Censo Agropecuário 1995/96 continua sendo o único¹ realizado e publicado no Brasil.

Dessa forma, dados censitários de tais atividades colaboram com futuras decisões/elaborações de políticas públicas conhecidas a realidades e peculiaridades das culturas desse imenso Brasil. Para o Estado do Rio Grande do Sul (RS), da mesma forma, por sua importância no fornecimento de alimentos para o Brasil e mercado externo torna-se fundamental conhecer o potencial de determinadas regiões.

Em estudo realizado², com dados do Censo Agropecuário 1995/96 somados a base de dados da Fundação de Economia e Estatística (FEE), o então Governo Germano Rigotto buscou amparo para medir o desempenho dos Conselhos Regionais de Desenvolvimento (COREDEs) do Estado do RS. Foi revelado um Rio Grande do Sul complexo. O Estado dividido pelos COREDEs, foi dividido conforme sua predominância de atividades, como: industrial; agrícola e mista. Assim como sua situação econômica, que varia entre COREDEs em dificuldades, estáveis, transição e dinâmicos.

Outrora, entre 1940 até meados dos anos 80, o RS era simplesmente tido como celeiro do Brasil, agora, com o deslocamento da fronteira agrícola para outras áreas do Brasil, o Estado se revela não mais dependente única e exclusivamente dos produtos agropecuários. O Rumos 2015 também revelou que os COREDEs de perfil agropecuário encontram-se com o nível desenvolvimento abaixo da média estadual. Ainda, mostrou que os COREDEs Fronteira Oeste, Jacuí Centro e Missões enfrentavam as maiores dificuldades. Em transição para uma matriz mista estariam os COREDEs Médio Alto Uruguai, Nordeste e Noroeste Colonial.

O Produto Interno Bruto (PIB) do Agronegócio, medido pelo Centro de Estudos Avançados em Economia Aplicada (CEPEA), no Brasil vem se recuperando e, atualmente, representa cerca de 20% do PIB nacional. Dessa forma, o PIB nacional não depende mais tão somente de produtos agrícolas ou do complexo agroindustrial. Diferentemente dos anos 30, onde as divisas das exportações eram em grande parte advindas do “grão” gerador de divisas,

¹ Em 1985 foi realizado um Censo Agrícola, mais restrito que o de 1995/96.

² Estudo encomendado pelo então Governador do Estado Germano Rigotto junto Fundação de Economia e Estatística denominado de Rumos 2015.

o café. Ainda, o Estado do RS, vive essa realidade dando espaço às novas fronteiras de produção que se estendem por Mato Grosso, Mato Grosso do Sul e à Amazônia.

No Estado, o complexo agropecuário representou nos últimos anos aproximadamente cerca de 40% do total das exportações gaúchas e, em 2006, 1,62% das exportações brasileiras. Ainda, nos últimos dez anos, em média, a agropecuária responde por 14,3% do PIB do Rio Grande do Sul. Não obstante, se torna latente a importância das atividades relacionadas à agropecuária para o Estado.

No Rumos 2015, a técnica da análise multivariada de dados foi utilizada para mapear os perfis dos COREDEs. Contudo, algumas questões importantes ficaram sem respostas. A própria metodologia não estava discutida e nem conclusões pormenorizadas foram expostas. Ainda, os COREDEs Alto Jacuí e Alto da Serra do Botucaraí que são intensivos em capital e especializados na monocultura da soja, da mesma forma, são regiões dinâmicas na produção de fumo, que é predominantemente familiar. Esta dicotomia dos COREDEs leva este estudo a buscar a relação entre os municípios que fazem parte dos COREDEs agrícolas. Ainda, existe a evidente hipótese de, em virtude desta dicotomia, se apresentar uma elevada diferença nos padrões e níveis de modernização do campo Gaúcho.

Deve-se ressaltar, tratando de um mundo globalizado, em que competitividade/eficiência e bons resultados caminham juntos, a indústria, a agricultura e pecuária não podem ficar atrás. Atualmente, para manter a competitividade e reduzir custos com mão-de-obra é cada vez mais comum – para não dizer indispensável – o uso de tecnologias avançadas, seja em sementes – biotecnologias –, maquinários e/ou técnicas de plantio.

É fato que a agricultura sofreu mudanças drásticas no Mundo todo. Dessa forma, o aumento de produção que até meados dos anos 50 foi baseado em aumentos da área plantada, sob governo dos militares, o padrão começa a se reestruturar. Pois, o país adotou uma nova estratégia de modernização, cujos reflexos no setor agrícola traduziram-se no modelo de complexos agroindustriais, reduzindo, assim, a agricultura a um subsetor da indústria. Pois, de um lado se encontra a indústria fornecedora de insumos modernos e maquinário e, de outro, as indústrias processadoras de alimentos e matérias-primas agrícolas.

Pese o fato que todo esse processo de mudança, principalmente na época do crédito agrícola farto, não tenha gerado os níveis de desenvolvimento almejados. Pois, muito deste crédito, ao invés de investimentos em tecnologia foi investido em especulações agrárias e, também, manteve um padrão conservador de modernização que se refletiu em todo o Brasil. Mesmo que os financiamentos atuais não tenham mais o apoio/subsídio do Governo, o

crescimento do produto da agropecuária continua se dando pelo aumento da produtividade da terra e aumento da utilização de insumos modernos resultados das pesquisas agrônomicas. Assim, a competitividade se instala como chave para a sobrevivência na atividade agropecuária.

Sendo assim, o presente trabalho se propõe a responder “em que medida, ou ainda, em que grau se encontrava a modernização agrícola nos COREDEs predominantemente agrícolas com base nos dados censitários coletados em 1995/96 pelo Instituto Brasileiro de Geografia e Estatística (IBGE)”? Dessa forma, esta resposta servirá de parâmetro comparativo com os dados do novo Censo Agropecuário 2007/08 e, também, como um alicerce para outros trabalhos do tipo no meio acadêmico e, até mesmo, para tomada de decisões políticas de futuros governantes eleitos no auxílio a regiões mais atrasadas ou mesmo deprimidas no Estado.

1.1 Objetivos

1.1.1 Objetivo geral

Inferir sobre o então estágio de modernização agrícola dos municípios nos COREDEs agrícolas do Rio Grande do Sul.

1.1.2 Objetivos específicos

- a) Pelo significativo número de variáveis – 17 – e municípios – 139 – objetiva-se formar grupos o mais homogêneos possível por meio da Análise de Agrupamentos;
- b) Identificar os fatores que representam com mais fidelidade a variação dos dados originais através da Análise Fatorial;
- c) Depois de agrupados e encontrado os fatores, destacar os padrões de modernização existentes nos grupos formados;
- d) Apresentar conclusões e tecer considerações acerca do grau de modernização agrícola gaúcha do período de 1995/96.

1.2 Justificativa

Tendo por base o Censo de 1995/96 torna-se possível inferir os níveis de modernização da agropecuária no Brasil, pois, este conta com dados referentes a número de tratores, máquinas para colheita, adubação, área explorada, população rural, investimentos, financiamentos, receitas, etc. Dessa forma, as informações acerca da agropecuária brasileira – do Rio Grande do Sul no caso –, aliado a modernos *softwares* e técnicas disponíveis permitem a avaliação dos níveis de modernização da agricultura gaúcha de maneira mais pormenorizada que o estudo Rumos 2015. Para isso, em virtude do elevado número de observações necessárias para estudo faz-se uso do instrumento estatístico da análise multivariada de dados quantitativos.

1.3 Estrutura do trabalho

O presente trabalho está dividido em seis seções, sendo a primeira esta introdução ao problema, seguida da metodologia escolhida, uma terceira e quarta que versam sobre a história da modernização da agricultura no Brasil e Rio Grande do Sul. Esta parte inclui a revisão da literatura sobre a análise multivariada, para finalmente, inferir sobre os resultados e culminar com a conclusão do estudo proposto.

2 METODOLOGIA E FONTE DE DADOS

A metodologia é um conjunto de procedimentos e técnicas utilizadas no processo de investigação. É fato, que a mesma é aplicada conforme a escola do pesquisador, ou mesmo, sua posição ideológica – por mais que este tente se eximir desta.

A prática da pesquisa consiste em um conjunto de técnicas que permitem o desenvolvimento da pesquisa nos diferentes momentos de seu processo. Pode-se denominar de pesquisa a busca de respostas/soluções para determinado problema, meras investigações ou a construção de um conjunto de conhecimentos que auxilia na compreensão da situação podendo orientar em ações futuras.

O desenvolvimento deste trabalho consistiu de pesquisas bibliográficas e levantamento de dados secundários possibilitando, assim, inferir sobre os objetivos propostos. Dessa forma, para desenvolver-se, este tomou um caráter teórico-empírico.

Seguindo a revisão de literatura, fez-se a tabulação eletrônica dos dados coletados e a análise dos mesmos por meio da aplicação de ferramentas da análise multivariada. Tal ferramenta foi crucial, pois, tornou viável o trabalho com um grande número de variáveis e observações, possibilitando a simplificação estrutural dos dados, sem comprometer a essência destes. A aplicação de técnicas multivariadas – Análise de Agrupamentos e Fatorial – permite identificar os padrões de modernização nos municípios gaúchos e, ainda, tornar possível hierarquiza-los os mesmos além de identificar as variáveis mais relevantes e possíveis potenciais de desenvolvimento.

Além do pacote Microsoft Office 2003 *for Windows* utilizaram-se os *softwares Statistical Package for Social Sciences*³ (SPSS) e *Statistica 7.0* disponibilizado pelo departamento de Estatística da Universidade Federal de Santa Maria – Centro de Ciências Naturais e Exatas.

Foram avaliados 139 municípios do Estado do RS sob a ótica de 17 variáveis. Os municípios fazem parte dos COREDEs essencialmente agrícolas⁴, que conforme classificação da Fundação de Economia e Estatística (FEE) foi adotado junto ao Governo do Estado do Rio Grande do Sul (2005) são: a) Fronteira Oeste, com 10 municípios; b) Jacuí Centro, com 6 municípios; c) Missões, 21 municípios; d) Noroeste Colonial, com 29 municípios; e) Médio

³ Versão *trial/demo* é encontrada em: <<http://www.spss.com/spss>>.

⁴ No ANEXO A encontra-se o mapa do Estado do Rio Grande do Sul e os COREDEs em questão.

Alto Uruguai, também, com 29 municípios, f) Nordeste, com 18 municípios, g) Alto da Serra do Botucarái, com 14 municípios e, h) Alto Jacuí, com 12 municípios.

As variáveis levantadas sobre os municípios dos referidos COREDEs foram divididas em três dimensões. Uma primeira é a das variáveis agropecuárias oriundas na totalidade do Censo Agropecuário 1995/96 que medem *stricto sensu*, os níveis de modernização da agricultura. Dentro deste grupo encontram-se variáveis que mensuram o número de estabelecimentos com assistência técnica especializada à agricultura e pecuária, número de tratores, máquinas para colheita, utilização de adubos, irrigação, os valores⁵ advindos da atividade agrícola e pecuária, receitas agropecuárias, os níveis/valores financiados, a área explorada e o efetivo bovino. Sendo objetivo analisar os possíveis e diferentes níveis de modernização nos municípios, as variáveis que se referem à pecuária foram utilizadas no intuito de isolar um fator para pecuária, contudo, como se observa adiante, o poder deste não teve sua raiz característica maior que 1.

Uma segunda dimensão socioeconômica, que é medida pela população rural e população economicamente ativa, observações estas oriundas também do Censo 1995/96 aliadas ao índice bloco renda⁶ dos municípios, calculados pela Fundação de Economia Estatística.

Por fim, especula-se uma questão espacial, com uma variável que mede a distância entre o município e o porto de Rio Grande, por onde escoia maior parte da safra exportada do RS. Considerando que, segundo Gama (2007, p. 27), no Rio Grande do Sul o transporte de cargas e mercadorias está baseado praticamente do modal rodoviário, ou seja, 90% do total do Estado. As variáveis, 17 no total, estão dispostas no Quadro 1 conforme sua sigla e fonte:

⁵ Os valores não foram deflacionados, pois, não se tratando de um comparativo em dois momentos, ou mesmo, de séries temporais, os resultados não se alteraram.

⁶ Este indicador medido pela FEE leva em consideração o Valor Adicionado Bruto (VAB) no comércio, alojamento e alimentação e Geração de Renda *per capita*. O presente trabalho versa sobre os COREDEs essencialmente agrícolas, dessa forma justificando a utilização do bloco renda, que aparentemente em sua medida, mede o todo do município, mas, como em tais COREDEs os reflexos da variação do bloco renda tem forte ligação com o campo – lembrando que são predominantemente agrícolas.

Variáveis agropecuárias		Descrição
01	ASSPEC	Assistência Técnica Especializada na Produção Animal/total de estabelecimentos no RS
02	ASSAGR	Assistência Técnica Especializada na Produção Vegetal/total de estabelecimentos no RS
03	TRATOR	Número de tratores/total de tratores no RS
04	MAQCOL	Números de máquinas para colheita/ total de máquinas para colheita no RS
05	ADUBOS	Número de estabelecimentos que utilizam adubos químicos/total de estabelecimentos no RS
06	IRRIGA	Número de estabelecimentos que utilizam sistemas de irrigação/total de estabelecimentos no RS
07	RECVEG	Valor das receitas advindas da venda de produtos vegetais/valor total no RS
08	RECANI	Valor das receitas advindas da venda de animais/valor total no RS
09	FINANC	Valor dos financiamentos nos municípios/valor total no RS
10	VALVEG	Valor da produção vegetal/valor total no RS
11	VALANI	Valor da produção animal de grande porte/valor total no RS
12	AREXPL	Área explorada nos municípios/total da área do RS
13	GADO	Efetivo de animais de grande porte do município/total do RS
Fonte: Todas as variáveis acima são oriundas do Censo Agropecuário do IBGE 1995/96, assim como, as variáveis PEA e POP que seguem.		
Variáveis Sócioeconômicas		Descrição
14	POP	População rural do município/total da população rural no RS
15	PEA	População economicamente ativa do município/total da população economicamente ativa no RS
16	RENDA	Índice bloco renda no município
A variável RENDA é um índice utilizado no Índice de Desenvolvimento Socioeconômico para o Rio Grande do Sul mensurado pela FEE. O Idese é um índice sintético, inspirado no IDH, que abrange um conjunto amplo de indicadores sociais e econômicos classificados em quatro blocos temáticos: Educação; Renda; Saneamento; Domicílios; e Saúde todos variando entre 0 e 1.		
Variável Espacial		Descrição
17	DISTRG	Distância do município em relação à cidade portuária de Rio Grande/total das distâncias em relação ao porto
Variável obtida junto ao Guia 4 Rodas.		

Quadro 1 – Variáveis que compõem o estudo.

3 UMA AGRICULTURA BRASILEIRA MODERNA – A ERA DAS MUDANÇAS ESTRUTURAIS

Esta seção, por abrir o item da revisão bibliográfica, tem o objetivo de apresentar a realidade vivida pela agricultura brasileira de meados do século XX até o presente momento. De forma sucinta apresentam-se os principais mecanismos de estímulo à modernização da agricultura nacional que viveu distintos momentos/fases que se alternaram conforme a realidade econômica por que passou o país.

3.1 Do Desenvolvimento e Modernização da Agricultura à Realidade Inflacionária

A idéia de um Brasil moderno, pelos militares, marcaria a economia nacional até a instauração do Plano Real, 30 anos após o golpe militar. A agricultura brasileira após o governo dos militares assumiria um novo rumo. Nem tão novo assim, pois ainda continua tendo como função gerar divisas, contudo, no referido momento da história, pretendia-se além de modernizar, diversificar a produção agrícola nacional.

Dessa forma, em meados dos anos 60, foi implantada no Estado o Programa de Melhoramentos da Fertilidade do Solo – Operação Tatu. Esta foi promovida por instituições de ensino – universidades –, pesquisa e extensão rural visando o aumento da produtividade das culturas via correção do solo gaúcho. Tal correção seria através da calagem do solo com o objetivo de favorecer o cultivo no Estado, principalmente, da cultura da soja.

Mas, a Operação Tatu serviu apenas como uma preparação para a alavancagem da produção agrícola no Brasil, principalmente, no Rio Grande do Sul (RS), por sua grande importância no setor de grãos e carnes que perdurou até o aparecimento das novas fronteiras agrícolas. Sabe-se que o período denominado de “milagre econômico” – final da década de 60 – promoveu o crescimento recorde da economia brasileira.

O Brasil, por sua “vocaç o agr cola”, teve a agricultura beneficiada pelo processo de moderniza o que acontecia no pa s, favorecida principalmente por uma pol tica de cr dito agr cola farto, para financiar n o s o as safras, mas garantir aumentos horizontais na  rea plantada e investimentos em tecnologia nas lavouras e infra-estrutura. Sendo que, no per odo de 1965-79, instaurou-se a fase de expans o no financiamento agr cola, mediante recursos do

governo, marcando o auge do Sistema Nacional de Crédito Rural (SNCR), onde fora oferecido abundantemente e com taxas de juros negativas para estimular a aquisição dos insumos considerados pelo governo como modernos e indispensáveis.

No período de 1950 a 1975, a economia crescia em média mais de 8% a.a. e a agricultura, seguindo o movimento, cresceu entre 4 e 5% ao ano em média, tendo a expansão da fronteira agrícola⁷ como principal responsável pelos aumentos da produção. Dessa forma, entre os anos de 1953 e 1966, os empréstimos do Banco do Brasil (BB), para aquisição de maquinário agrícola, mais que quadruplicaram em termos reais e, os empréstimos triplicaram (BACHA, 2004).

Contudo, esse período de euforia foi alavancado sob fortes penas, repercutindo no alargamento do déficit público e o aumento da pressão inflacionária. Conforme Gremaud et al. (2002, p. 410), as conseqüências dessa opção, bem como das alterações no cenário internacional no final da década – novo choque do petróleo em 1979, alteração substancial da política econômica norte-americana que leva à moratória Mexicana⁸ – sentenciaram a década seguinte, período marcado pela recessão da economia brasileira em função da reação à chamada “crise da dívida externa” contraída na vigência do então governo militar.

Segundo Graziano da Silva (1996, p. 107), “no final dos anos 70, os indícios de uma crise já eram evidentes”. A inflação havia superado o patamar de 40% ao mês, as taxas de crescimento do Produto Interno Bruto (PIB) declinavam ano a ano, a dívida externa já alcançava os 60 bilhões de dólares junto a uma taxa de desemprego crescente.

Seguindo a sucessão de planos econômicos, o período entre os planos Cruzado e Collor II marca um período em que o combate da inflação vinha como prioridade. Dessa forma, implicando mudanças para o crédito rural, que no período anterior foi o maior responsável pelo processo de modernização da agricultura nacional, possibilitando aumentos reais da produção. Posteriormente, o Plano Real também teve forte influência para a agricultura em sua implantação e vigência.

⁷ Alves & Pastore (1975, p. 397 apud, SANTOS, 1998) complementam a idéia: que no período entre 1950-70, 70% do aumento da produção agrícola era explicado pela expansão da área cultivada e 30% devido ao aumento da produtividade da terra.

⁸ Prorrogação de prazo solicitado pelo devedor, ou concedido pelo credor, para pagamento de uma dívida. Segundo Freitas (2003), sob a presidência de Paul Volcker (1979 a 1987), nomeado pelo Presidente Jimmy Carter, o FED elevou os juros de 3,5% para 10% ao ano, a fim de conter as pressões inflacionárias, provocadas pela crise do petróleo (1979). A América Latina, principalmente, sofreu com as tensões provocadas pelo arrocho inflacionário da política monetária dos Estados Unidos, com a disparada do preço do dólar e a retração de oferta do crédito externo. Com isso, em agosto de 1982 a economia Mexicana veio a ruir e decretar moratória.

3.1.1 Os Planos Heterodoxos⁹: O comportamento da intervenção do Estado na agricultura - 1986 a 1994

O período que compreende o ano de 1985 até o início da década de 90, foi marcado pela preocupação no combate à inflação, e assim, resultando na dificuldade do Estado em dar seguimento à política expansiva dos governos anteriores à (re)democratização. O governo mudou, as políticas de gastos públicos seriam refeitas e a agricultura sentiria os reflexos – o que veio por caracterizar o “período das dúvidas” para a cultura da soja e para o Estado, periodização que fora ressaltada por Brum (2004).

Outra marca da década, além da crise que se avizinhava, era a transição de governo do regime militar para um regime democrático. O Presidente Tancredo Neves, eleito pelo colégio eleitoral, nem chegou a assumir o cargo vindo a falecer em 21 de abril de 1985. Em 22 de abril, o vice José Sarney foi nomeado oficialmente no cargo de presidente do Brasil. Do ponto de vista econômico, o governo Sarney foi bastante conturbado. Herdeiro de problemas anteriores de desenvolvimento gerados durante o antigo regime militar e agravados pelas crises internacionais, o governo elaborou seu primeiro plano de combate à inflação em 28 de fevereiro de 1986.

Com o Plano Cruzado e uma série de medidas heterodoxas, o Brasil teve mais um falecimento de moeda. O cruzeiro foi substituído pelo cruzado, junto a um corte de três zeros. Em um curto espaço de tempo, a inflação foi reduzida, o desemprego diminuiu e o poder aquisitivo da população aumentou. Mas em poucos meses o Plano Cruzado começou a demonstrar sintomas de crise.

Logo, houve o desaparecimento de produtos do mercado, e a cobrança de ágio, ou seja, o consumidor era obrigado a pagar uma parcela a mais sobre o preço estipulado pelo congelamento. Em novembro de 1986, veio o Plano Cruzado II, que também congelou os preços, porém, acima da realidade de mercado. Com o passar do tempo, a inflação voltou a crescer e, em maio de 1987, já ultrapassava o percentual dos 20% ao mês elevando o preço dos combustíveis (60,16%) e dos automóveis (80%). Este evidente fracasso do plano junto a perda da confiança da população com o governo veio pressionar a substituição do Ministro Dílson Funaro.

⁹ Segundo Gremaud et al. (2002, p. 444), para os heterodoxos a inflação não decorre de excesso de demanda provocado pela emissão monetária. A emissão monetária é vista muito mais como uma decorrência da inflação do que como uma causa. A característica desta linha heterodoxa de economistas é adotar medidas de congelamento de preços, salários e controle do câmbio como poderá ser observado no decorrer da seção.

Luiz C. Bresser Pereira, no dia 29 de abril de 1987 assumia o poder do Ministério da Fazenda junto aos sérios problemas da inflação – um mês depois de sua posse a inflação alcançou os 23,26% ao mês.

O Plano Bresser foi apresentado em junho de 1987, quando foi decretado um congelamento de preços, aluguéis e salários por um período de 60 dias. E para diminuir o déficit público algumas medidas foram tomadas, como: aumentou tributos, eliminou o subsídio do trigo e adiaram-se grandes investimentos em obras de infra-estrutura que estavam planejados dentre outros. Mesmo após todas estas medidas adotadas a inflação continuou sua rigidez. A situação crônica leva mais uma vez a troca do poder no Ministério da Fazenda (MF), quando assume este ministério Maílson da Nóbrega em 6 de janeiro de 1988.

Maílson assume o poder com um discurso de uma política econômica sem medidas drásticas, tentando corrigir a inflação com ajustes localizados, denominando estas medidas como um “Feijão com Arroz”, na qual deveria evitar a hiperinflação. Porém, em 1988, a inflação chega a 993% e, por conseguinte, outro plano falhara e fora decretado um novo.

O Plano Verão foi apresentado por Maílson da Nóbrega em 15 de janeiro de 1989. Novamente aconteceu um corte de três zeros e criação de uma nova moeda, o cruzado novo. Foi extinta a correção monetária, proposto a privatização de algumas estatais e novamente, um enxugamento dos gastos públicos, onde os funcionários públicos contratados nos últimos cinco anos seriam exonerados. Os cortes não foram feitos, o plano fracassou e a inflação disparou.

Só para se ter idéia da gravidade em que se encontrava a economia, de fevereiro de 1989 até fevereiro de 1990, a inflação ultrapassava a casa dos 2.000% no período.

As dificuldades não se restringiram somente ao povo, com constantes perdas no padrão de consumo, também atingiram o setor agrícola, com a incapacidade agora reconhecida do governo de continuar a abastecer a agricultura com subsídios e crédito farto. Marcando assim segundo Graziano da Silva (1996), “[...] um novo padrão de desenvolvimento baseado no crescimento intensivo via aumentos dos rendimentos físicos” e não mais no constante aumento das áreas cultivadas¹⁰. Essa busca de rendimentos se faz diante de aumentos da produtividade da terra, sendo que devido às dificuldades econômicas do período seria a principal maneira para driblar os problemas.

¹⁰ Santos (1998, p. 395-6) ratifica a idéia, citando que a ênfase das políticas de crédito por parte do governo, que foram sempre na medida de procurar aumento no produto agrícola se deu via expansão da área utilizada. Interessante é o fato de que, no período, a marca sempre foram os baixos níveis alcançados de produtividade pela agricultura nacional.

A crise em que vivia o país nos anos 80, implicou uma nova orientação dos gastos públicos voltados para a função agricultura. Esse momento ficaria marcado pela escassez de recursos para o financiamento rural e um maior destaque para as políticas de preços mínimos¹¹. Barros & Araújo (1995, p. 01) salientam que, a política rural foi alterada em dois sentidos: primeiro, o montante real de recursos sob disposição da agricultura foi se reduzindo quase que continuamente e, segundo, esforços foram efetivados objetivando diminuir os subsídios implícitos nas taxas de juros oferecidas.

Dessa forma, a passagem dos anos 80 para a década seguinte consolidou um período extremamente conturbado para a economia e, conseqüentemente, para o agronegócio. O que se esperava deste quadro negativo para a agricultura, era que seu crescimento fosse interrompido.

Segundo Graziano da Silva (1996) a agricultura nesse período fora o único setor da economia que conseguiu elevar sua produtividade ao longo dos anos 80, embora isso não tenha garantido um processo de acúmulo de capital para o setor durante este período, o que caracterizaria uma *década perversa*¹² para a agricultura. Dias (1995) corrobora mesma idéia em seus estudos, concluindo que o setor agrícola teve um desenvolvimento favorável, mesmo considerando a crise que a economia enfrentava.

A redução de crédito rural provocou impactos diretos, porém, de curto prazo na produção, como a redução na utilização direta de fertilizantes e um encolhimento na área colhida.

Não tardando, uma recuperação foi possível com o aumento de produtividade mediante a utilização de fontes internas de capitalização ou de crédito alternativo suficientes para este processo de ajustamento com a nova realidade (DIAS, 1995, p. 16).

A falta do vigor econômico no período dos anos 80 levou a uma redução do ritmo da modernização da agricultura no país. Segundo Brum (1984), isto se deveu a fatores como: redução dos incentivos do governo, como crédito, para o processo de modernização e

¹¹ Política Governamental que consistia no estabelecimento de preços pelo governo, com os quais poderiam ser comprados determinados produtos caso os produtores não encontrassem melhores alternativas de comercialização no mercado. O governo compraria os excedentes do mercado, formando os chamados estoques reguladores de preços, para então revendê-los nas épocas de entressafra, com a finalidade de evitar uma alta demasiada nos preços dos produtos agrícolas.

¹² Período fortemente recessivo da economia brasileira, que vinha de um cenário de forte crescimento econômico dos anos 70, o qual o país se depara com quedas expressivas no Produto Interno Bruto nos anos de 1981 e 1983. Mas segundo Graziano da Silva (1996, p. 109-110), para a agricultura ficou caracterizado o período não como Década Perdida, mas sim como: *Década Perversa*. Sendo que as perdas não foram tão expressivas se comparado à economia do país, tendo em vista que a agricultura ainda contribuiu para que o PIB global não tivesse quedas ainda mais expressivas.

aumentos horizontais; o caráter desigual e excludente do processo de modernização agrícola, fazendo com que produtores e regiões mais “aptas” à modernização estivessem inseridos nesse processo – como é o caso do Rio Grande do Sul entre os anos de 1970/80 – e a mudança do padrão de modernização da agricultura, baseado na incorporação de “novas tecnologias” – informática, microeletrônica e biotecnologias¹³.

Algumas características do caráter desigual e excludente¹⁴ que podem ser consideradas são as baixas taxas de juros praticadas para o empréstimo rural, influenciando os agentes financeiros, que tinham como preferência financiar aqueles produtores com maiores garantias, liquidez e maior transferência de recursos para determinada atividade.

Segundo Fagundes (1986), essa é a forma de identificar os subsídios implícitos no financiamento do setor rural. Os grandes produtores se tornaram os principais favorecidos. Isso prova que a concentração do crédito ocorreu nas grandes propriedades e principalmente nas culturas de exportação – como a soja na região do RS (BACHA, 2004, p. 65).

Sayad (1984 apud MASSUQUETTI, 1998, p. 106) afirmou que os empréstimos com juros subsidiados “representa gasto distribuído de maneira desigual, visto que se atribuiu em proporção maior para os maiores produtores” e pressionam as contas do Governo, a base monetária, a liquidez da economia e, conseqüentemente, a inflação.

A evolução dos subsídios implícitos no crédito rural, durante o período 1969/86, pode ser vista no Tabela 1, onde se verificam as taxas de juros reais negativas subsidiadas praticadas no auge do crédito rural, evidenciando assim uma pressão nas contas nacionais, visto os juros.

¹³ Embora o papel da biotecnologia nos anos 80 tenha uma atuação dita ainda modesta, e só vindo a eclodir depois de 2000.

¹⁴ Bacha (2004, p. 65) ratifica essa idéia de desigualdades no processo de modernização onde “o crédito rural beneficiou, principalmente, as culturas de exportação, as regiões Sul e Sudeste (zonas antigas de produção) e os médios e grandes produtores rurais”. Ainda, Santos (1998, p.401) corrobora com a idéia de Bacha, ao afirmar os principais fatores componentes de crédito agrícola: a) taxa de juros b) risco da atividade em que o crédito vai ser aplicado, c) risco associado à firma tomadora de empréstimo e d) garantias oferecidas ao credor; onde o risco menor seria um empréstimo aos grandes produtores, devido a uma maior capitalização.

Tabela 1 – Taxa real de juros do crédito rural no Brasil – 1969/86.

ANO	TAXA REAL DE JUROS
1969	-1,80
1970	-1,90
1971	-2,30
1972	-0,40
1973	-0,60
1974	-16,00
1975	-12,00
1976	-22,70
1977	-17,80
1978	-19,00
1979	-36,20
1980	-41,80
1981	-26,60
1982	-25,80
1983	-26,30
1984	0,90
1985	-1,80
1986	-36,80

Fonte: Caffagni (1995, p. 27).

Os aumentos de produção referentes à década de 80 seriam então baseados em um novo padrão: a produtividade. Até a década de 70, o crescimento da agricultura foi baseado em função da expansão horizontal¹⁵ do plantio e de mão-de-obra empregada, estando o processo de ganhos gerados a partir de produtividade num estágio ainda prematuro.

Este aumento da produtividade nos anos 80, levando em consideração a crise do período e os preços desfavoráveis¹⁶ para a atividade, segundo Gasques & Villa Verde (1990, 1997) pode ser explicado pelos seguintes argumentos: o papel das pesquisas agronômicas influenciando na produtividade, ressaltando o papel ativo da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA); até então o aumento na utilização de determinados insumos, em função da redução de custos e a mudança na composição da produção, ou seja, um deslocamento para a produção de produtos com maior valor comercial e mercado.

Quanto ao aspecto das políticas governamentais como fator de explicação para o aumento de produção, Gasques & Villa Verde (1990, 1997) salientam alguns dos principais efeitos:

- a) com uma política mais restritiva de crédito, 70% do total de recursos deveriam ser aplicados como crédito de custeio de safra;

¹⁵ Aumentos de produção, levando em consideração a área plantada, onde os aumentos de produção obtidos estão relacionados a inclusão de novas áreas para cultivo, ou seja, uma utilização maior de recursos em lugar de uma melhor utilização dos recursos.

¹⁶ Monteiro (1999) em seus estudos onde comparou os preços dos produtos agrícolas do período 1990/98 com o mesmo da década anterior (1980/89). Percebe-se uma redução de 52,45% pelo preço a ser recebido da soja no Brasil, contribuindo para com as observações de Gasques e Villa Verde.

- b) a Política de Garantia de Preços Mínimos (PGPM)¹⁷ sofreu alterações que a transformaram no principal instrumento da política agrícola do período, atuando com um redutor de riscos na comercialização dos produtos colhidos, limitando a atuação do governo no financiamento do setor agrícola, onde os recursos seriam administrados caso a caso;
- c) o Programa de Garantia da Atividade Agropecuária (PROAGRO)¹⁸ agiu como um fator de redução dos riscos;
- d) a unificação dos preços mínimos (1981) estimulou a produção na fronteira;
- e) a relação positiva entre preços mínimos e custos operacionais garantiu o pagamento dos custos variáveis e permitiu a obtenção de uma margem de lucro;
- f) a uniformização do preço do óleo diesel;
- g) os programas, de desenvolvimento regional e de incentivos fiscais também influenciaram na expansão da produção.

Para Gasques & Villa Verde (1990, p. 15) as características da agricultura como os aumentos de produtividade, as quedas nos custos de produção juntamente com as políticas agrícolas, foram os principais fatores que possibilitaram o crescimento da agricultura, mesmo com preços reais dos produtos em constante decréscimo junto às adversidades no período.

Silva (1996 apud MASSUQUETTI, 1998, p. 112) acrescenta que

se praticou, do lado da produção agrícola, uma política agressiva de preços e contencionista de crédito, logrando-se a proteção da rentabilidade para determinados segmentos empresariais. Tal política se expressou em: concessões de crédito de custeio a taxas reais de juros negativos, mantendo-se a oferta de fundos relativamente estável ou ligeiramente declinante entre 1980 e 1984 e uma queda acentuada na concessão de crédito para investimento e elevação das taxas nominais de juros.

A crise do setor agrícola brasileiro e gaúcho no final dos anos 80 foi uma das conseqüências do processo inflacionário vivido no período, que expôs as dificuldades enfrentadas pelos produtores rurais com o agravamento da inflação no período posterior ao Plano Cruzado. Aliada a essa pressão inflacionária da década, está a forma como o então governo veio por conduzir a política econômica em busca da estabilização, e os próprios efeitos que

¹⁷ A PGPM conforme Bacha (2004, p. 67) em 1943 foi criada a Comissão de Financiamento da Produção (CFP) – atual Companhia Nacional de Abastecimento (CONAB). Junto a CFP foi instituída a Política de Preços Mínimos, onde os primeiros preços mínimos foram fixados em 1945, referindo-se a produtos como a soja, feijão e milho. Em 1990 a PGPM abrangia 42 produtos excluídos café e cana-de-açúcar, que possuíam políticas próprias.

¹⁸ Segundo Bacha (2004, p. 77), o PROAGRO tinha a finalidade de exonerar o produtor rural do cumprimento de obrigações financeiras concernentes ao crédito rural caso ocorressem adversidades climáticas que inviabilizassem o pagamento da dívida bancária. Tratava-se de uma forma de seguro do crédito rural. Se por motivos naturais, pragas ou doenças o produtor perdesse sua safra, o PROAGRO cobriria o crédito rural devido.

essas políticas de estabilização vieram a exercer sobre o setor agrícola brasileiro. Rezende (1990, p. 510 apud GRAZIANO DA SILVA, 1996) conclui que

[...] o setor agrícola brasileiro foi profunda e adversamente afetado pelo agravamento do quadro macroeconômico e, em particular, pelas estratégias – seja via choques seja via moeda indexada – adotadas pelo Governo para controlar a inflação. O setor agrícola que encontrou espaço e recebeu estímulos para crescer na maior parte da década de 80, tornou-se, ao término desta, a principal vítima do descontrole inflacionário e da incapacidade demonstrada pelo Governo de combatê-lo.

Até mesmo os produtos que possuíam políticas específicas – que não é o caso da soja – como trigo, o açúcar e o álcool, foram atingidos pela redução dos subsídios ocorrida no período entre 1988 e 1989 (MASSUQUETTI, 1998). Implícito fica o fato de a política de crédito rural nos anos 80 ter sofrido diversas alterações. No que tange à Política de Garantia de Preços Mínimos, Barros & Araújo (1995, p. 33-4) relatam que

[...] na década de 80, porém, com as mudanças havidas no crédito rural e as condições desfavoráveis no mercado externo, procurou-se ativar a política de preços, procedendo-se a aperfeiçoamentos na sua sistemática, como forma de compensar o setor agrícola e, assim, assegurar a manutenção de seu crescimento. Todavia, duas dificuldades levaram as autoridades a reduzir a ênfase nessa política. Por um lado, os estoques públicos tenderam a se acumular e seus custos passaram a onerar o orçamento além do previsto. Por outro lado, o acúmulo de estoques nas mãos do governo passou a aumentar os riscos de mercado, tendo se observado um afastamento do setor privado da comercialização na safra e das atividades de armazenamento. Parte desses problemas poderá ser atenuada com a recente regionalização dos preços mínimos. O estabelecimento de regras para a atuação do governo nos mercados agrícolas, e a estrita observância das mesmas são também muito importantes para atenuar os riscos mencionados. Entretanto, é fundamental, um elevado grau de realismo no tocante à definição do grau de estabilidade a ser alcançado com a política de estoques. Embora a estabilidade seja altamente desejável, seu custo para os cofres públicos deve ser devidamente considerado.

Com essa redução dos recursos oriundos do Tesouro Nacional e dos subsídios para o financiamento rural, a PGPM ficou mesmo como o instrumento mais utilizado pelo Governo, para assegurar os preços a serem recebidos pelos produtores rurais para garantir uma rentabilidade mínima do setor.

As reduções no volume de recursos, para o financiamento rural, podem ser observados no Tabela 2, que mostra a variação anual do volume de recursos para o crédito agrícola no período de 1969/94. Implícito está o volume crescente de crédito oferecido até 1979, quando atinge o ápice e também marcaria o aumento constante da inflação. Após este ano, ocorre uma queda contínua até 1984 e um ensaio de uma retomada no início do Plano Cruzado “com circunstâncias favoráveis criadas pela estabilização alcançada nos primeiros meses do Plano” (GRAZIANO DA SILVA, 1996, p. 115).

Tabela 2 – Inflação e a evolução dos recursos dos financiamentos concedidos aos produtores e cooperativas no Brasil – 1969/94.

Ano	Inflação	Valores Constantes em Reais (*)	Variação Anual (%)
1969	19,31	7.494.663.886	-
1970	19,27	8.954.766.759	19,48
1971	19,48	10.347.952.571	15,56
1972	15,73	12.802.480.059	23,72
1973	15,53	18.111.842.170	41,47
1974	34,56	22.391.525.971	23,03
1975	29,33	32.647.469.139	45,80
1976	46,27	33.450.757.289	2,46
1977	38,79	29.865.926.724	(10,72)
1978	40,81	30.365.914.847	1,67
1979	77,24	37.843.151.359	24,62
1980	110,23	36.190.757.419	(4,37)
1981	95,20	31.390.465.434	(13,26)
1982	99,73	30.397.617.678	(3,16)
1983	211,02	22.946.277.749	(24,51)
1984	223,81	14.015.813.900	(38,92)
1985	235,13	19.986.714.977	42,60
1986	65,04	29.803.712.681	49,12
1987	415,95	23.494.328.088	(21,17)
1988	1.037,53	16.578.837.029	(29,43)
1989	1.782,85	15.137.757.930	(8,69)
1990	1.476,71	8.649.116.329	(42,86)
1991	480,23	8.923.435.441	3,17
1992	1.157,84	9.894.149.647	10,88
1993	2.708,17	8.496.887.847	14,12
1994	1.093,85	12.275.946.328	44,48

Fonte: Anuário Estatístico do Crédito Rural – 1997.

Nota: (*) A preços de 1997 – Deflator: Média Anual do IGP-DI – FGV.

Os governos vigentes ainda procuravam investir contra a inflação com planos de choques heterodoxos. No governo Collor/Itamar Franco, a preocupação como não podia ser diferente, continuou sendo, também, o combate à inflação, atacando com um bloqueio de liquidez aliado a reforma monetária.

Em março de 1990, o Plano Collor, é apresentado para a sociedade. Acompanhando o mesmo, estavam: a substituição do cruzado novo pelo cruzeiro, um bloqueio de 18 meses dos saldos das contas correntes, cadernetas de poupança e demais investimentos superiores a Cr\$ 50.000,00. Os preços foram tabelados e depois liberados gradativamente. O plano também previa a abertura do mercado interno, com a redução gradativa das alíquotas de importação – o que marcaria a característica neoliberal do Governo Collor, assim como, um risco para alguns setores da agropecuária gaúcha acostumada a ser protegida pelo governo. Proteção esta que traz distorções da realidade competitiva de um setor.

O Plano Collor não diferia muito, no critério crédito rural, que fora utilizado pelo Plano Cruzado anteriormente. A atenção para com a participação do Governo no financiamento da agricultura foi novamente uma meta no Governo. E junto a todas estas medidas, mais cortes nos gastos públicos e demissões de funcionários junto a um processo de privatizações de empresas estatais.

Segundo Massuquetti (1998) no ano das reformas do Governo Collor, em 1990 em relação ao ano de 1989, os gastos do setor público da União cresceram 3,4%, enquanto os gastos com a agricultura reduziram-se em 50%.

Uma das medidas utilizadas neste novo Governo, o confisco dos ativos financeiros, teve um efeito negativo imediato sobre a maioria dos produtores rurais, visto que, os mesmos ficaram sem recursos para desenvolver a colheita dos produtos. Outra medida que prejudicou os produtores foi a correção dos financiamentos rurais pela variação do Índice de Preços ao Consumidor (IPC) de março (84%), sensivelmente superior à correção dos preços mínimos pela variação do Bônus do Tesouro Nacional (BTN) de março (42%), causando a impossibilidade para muitos produtores de pagamento das dívidas. Reflexos desses fatos podem ser observados na produção gaúcha de soja, que foi atingida em cheio pelos efeitos do plano, aliado os problemas climáticos – *La Niña*¹⁹ –, onde na safra de 90/91 colheu-se 2.220.502 toneladas e ainda, com uma redução de aproximadamente 5% na área colhida em relação ao ano anterior.

O processo de liberalização do mercado agrícola tinha o objetivo, segundo a política do Governo, de reduzir o seu papel e, conseqüentemente, os recursos do Tesouro no financiamento da atividade e tornar a agricultura brasileira mais competitiva no mercado mundial.

Outras medidas também foram adotadas com objetivo de uma redução da participação da União no setor rural, como a extinção do Instituto do Açúcar e do Alcool, Instituto Brasileiro do Café, Empresa Brasileira de Extensão Rural (EMBRATER) – que fora criada junto a EMBRAPA no Governo militar – e Banco Nacional de Crédito Cooperativo (BNCC) e; a partir da união da Comissão de Financiamento da Produção, da Companhia Brasileira de

¹⁹ O fenômeno *La Niña* consiste da ocorrência de águas mais frias do que a média na superfície do Oceano Pacífico equatorial. Esse fenômeno determina no Rio Grande do Sul estiagem, principalmente, no início do verão do ano de seu início, com "repique" no final do outono e início do inverno do ano seguinte. No período de maior efeito do fenômeno (outubro, novembro e dezembro) os estudos climatológicos, com séries históricas de 86 anos de observações, mostram que em cerca de 75% dos casos de ocorrência de *La Niña* as chuvas ficam muito abaixo da média climatológica. Em outubro e novembro, os dois meses mais críticos, a probabilidade de se ter chuvas muito abaixo da normal é maior do que 90% (EMATER, 2007).

Armazenamento (CIBRAZEM) e da Companhia Brasileira de Alimentos (COBAL) criou-se a Companhia Nacional de Abastecimento (CONAB)²⁰.

Massuquetti (1998) complementa que a política adotada pelo Governo Collor provocou uma série de conseqüências desastrosas para o setor agrícola na safra 1989/90, entre elas destacam-se a queda na taxa real de crescimento do PIB agropecuário (-4,4%); queda no rendimento físico de importantes produtos como o arroz, o feijão, a laranja, a soja, o milho e o trigo; e a redução do comércio de determinados insumos e máquinas – tratores, fertilizantes e calcário. Como resultado dessa política de redução da intervenção do Governo na agricultura e de desmonte do aparelho estatal voltado para a condução das políticas agrícolas, houve uma redução na quantidade colhida na safra de 1990/91 no Brasil e no Estado do Rio Grande do Sul.

Todo esse choque na economia, com a retirada de grande parte do dinheiro em circulação, deu início a uma grande recessão no país. As medidas adotadas acarretaram um aumento do desemprego, grandes empresas estavam sem liquidez para investimentos e muitas pequenas empresas acabaram por fechar, diminuindo a produção brasileira. Mais uma vez fica evidente a necessidade de um novo plano na tentativa de acalmar a economia.

O Plano Collor II é lançado em janeiro de 1991. Tinha como objetivo controlar a circulação financeira, extingue as operações de overnight²¹ e cria o Fundo de Aplicações Financeiras (FAF) onde centraliza todas as operações de curto prazo; acaba com o Bônus do Tesouro Nacional Fiscal, o qual era usado pelo mercado para indexar preços; passa a utilizar a Taxa Referencial Diária com juros prefixados e aumenta o Imposto sobre Operações Financeiras (IOF); pratica uma política de juros altos e faz um grande esforço para desindexar a economia e tenta mais um congelamento de preços e salários. O governo acreditava que aumentando a concorrência no setor industrial conseguiria segurar a inflação, então se cria um cronograma de redução das tarifas de importação, reduzindo a inflação de 1991 para 480%.

²⁰ A CONAB (CONAB, 1998) é uma empresa pública vinculada ao Ministério da Agricultura e Abastecimento que foi criada por Decreto Presidencial e autorizada pela Lei nº 8.029, de 12/04/90, iniciando a sua atividade em 01.01.91. A CONAB tem por finalidade executar a Política Agrícola no segmento do abastecimento alimentar e a Política de Garantia de Preços Mínimos, fornecer subsídios ao Ministério da Agricultura e do Abastecimento na formulação e acompanhamento das referidas políticas, bem como na fixação dos volumes mínimos dos estoques reguladores e estratégicos. A CONAB tem por objetivos básicos: I. - garantir ao pequeno e médio produtor os preços mínimos e a armazenagem para guarda e conservação de seus produtos; II. - suprir carências alimentares em áreas desassistidas ou não suficientemente atendidas pela iniciativa privada; III. - fomentar o consumo dos produtos básicos e necessários à dieta alimentar das populações carentes; IV. - formar estoques reguladores e estratégicos, objetivando absorver excedentes e corrigir desequilíbrios decorrentes de manobras especulativas; V. - participar da formulação da Política Agrícola; VI. - fomentar, por meio de intercâmbio com universidades, centros de pesquisas e organismos internacionais, a formação e o aperfeiçoamento de pessoal especializado em atividades relativas ao setor de abastecimento.

²¹ Operações realizadas no *Open Market* por prazo mínimo de um dia. No *overnight*, o saque estava limitado a 25 mil cruzeiros ou a 20% do saldo total.

Se o Governo Sarney ficou marcado pelo início incessante do combate contra a inflação, o Governo Collor foi responsável pelo início do processo de abertura econômica brasileira, com um programa de liberalização financeira externa e de eliminação de barreiras protecionistas contra a importação. O programa de eliminação de barreiras consistiu em: a) eliminação ou redução da cobertura de barreiras não-tarifárias, tais como reservas de mercado, quotas, proibições, etc; b) diminuição no nível médio das tarifas de importação; e, c) redução do grau de dispersão na estrutura tarifária.

O programa de redução tarifária estabeleceu a redução da tarifa média de 32,2%, no ano de 1990, para 14,2%, em 1994. Com a liberalização financeira externa, por sua vez, a economia brasileira integrou-se aos fluxos de capitais em busca de fontes de aplicações rentáveis em países em desenvolvimento, inclusive, na América Latina.

A economia começa a demonstrar sinais de recuperação no final de 1992, após um grande processo de reestruturação interna das indústrias brasileiras. Pode-se afirmar que fora fundamental a abertura do mercado brasileiro para a entrada de produtos importados, que obrigou a indústria nacional a elevar os investimentos na modernização do processo produtivo, qualidade e lançamento de novos produtos no mercado interno. Essas mudanças eram necessárias para as empresas se tornarem mais competitivas, tanto no mercado interno quanto no mercado externo. O aumento de produtividade foi fundamental para a sobrevivência das empresas e da agricultura.

Quanto aos instrumentos de política agrícola mais utilizados nesse período, destacam-se: a) Política de Garantia de Preços Mínimos (PGPM)²², onde têm-se, como principais medidas, o Valor Básico de Custeio (VBC) para o financiamento do custeio, o Preço Mínimo como garantia contra incertezas de preços, a definição de limite de Preço para a Liberação de Estoques públicos (PLE), os Empréstimos do Governo Federal (EGF) para a comercialização e EGF Especial, a Aquisição do Governo Federal (AGF), o Prêmio de Liquidação e a Equivalência Produto; b) PROAGRO, que busca reduzir o risco na atividade rural; c) financiamento para investimentos, cujo objetivo é melhorar a qualidade dos produtos agrícolas e aumentar a produtividade e d) apoio tecnológico.

Segundo Massuquetti (1998, p. 133), os instrumentos VBC, EGF e Preços Mínimos combinam “com uma política de economia aberta, mas também podem conter subsídios e outras formas de intervenção governamental” e o Prêmio de Liquidação e o EGF Especial

²² Ocorre em 1991, a regionalização dos preços mínimos, com intuito de incentivar a interiorização da agroindústria e desestimular atividades distantes de seus respectivos mercados. Segundo Alves & Paulillo (1998, p. 22), a PGPM executada no governo Collor, foi mais um apêndice da política de crédito rural, assim, caracterizando uma adaptação a conjuntura econômica do país.

“visavam transferir para a iniciativa privada ou pelo menos retardar a intervenção do Governo na comercialização de produtos garantidos pelos preços mínimos”. Portanto, estando mais como uma política de abertura da economia e o contínuo e gradual afastamento do governo, o que viria a caracterizar a agricultura e sendo reforçado pelo Plano Real futuramente.

No período entre 1990 e 1993, conforme o Banco Central (BC, 2004) existe uma grande, se não a maior, e mais prejudicial crise do setor agrícola, onde os valores de créditos concedidos pelo Sistema Nacional de Crédito sofreram uma grande redução: de R\$ 7,94 bilhões para R\$ 4,71 bilhões, no período. Os gastos governamentais com a agricultura representaram, no ano de 1993, 1,9% do dispêndio total do Governo – Balanço Geral da União, 2004. Os percentuais de 1990, 1991 e 1992 foram 1,9%, 4,4% e 2,5%, respectivamente. Foram aplicados, em 1993, aproximadamente 4,3 bilhões de dólares na agricultura, sendo que 3,6 bilhões foram destinados aos programas de Administração, Organização Agrária, Abastecimento e Recursos Hídricos (GASQUES & VILLA VERDE, 1995), ratificando a idéia do afastamento por parte do Governo relativo aos subsídios e crédito farto ativos nas décadas passadas, voltados para aumentos de produção.

Contudo, a implantação de diversos planos, conturbados e ineficientes até então para a economia brasileira, serviram para chamar a atenção da população e do mercado para com a credibilidade ou incapacidade de comprometimento das autoridades econômicas brasileiras com uma política anti-inflacionária efetiva e consistente. Certo é afirmar que a confiabilidade ao final de cada um dos planos de estabilização estava baixa. Com isso, a credibilidade governamental tem se tornado uma variável macroeconômica de grande relevância nos últimos anos. Fato esse que levou o então Presidente Fernando Collor de Melo ao processo de impeachment sendo substituído por Itamar Franco²³.

O início da década de 90 marca o esgotamento do modelo que foi usado durante os vinte anos anteriores, onde a hiperinflação era reprimida quase que em todos os planos anteriores, pelo congelamento de preços. Tamanha era a inflação que fora também decretado a falência da moeda por inflação.

²³ Impeachment de Fernando Collor - O Presidente da República foi substituído sem derramamento de sangue, golpe militar ou qualquer tipo de violência. Foi um processo pela via legal e demonstrou amadurecimento do povo e dos políticos brasileiros, o que foi excepcional para a América Latina. Collor pregava a moralidade, combate à corrupção, porém na gestão de seu governo foram constatados muitos casos de corrupção. Paulo César Farias, o PC Farias, braço direito e tesoureiro do Presidente Collor, foi relacionado no esquema de corrupção dentro do governo, sendo que a CPI apurou que muito dinheiro foi desviado para a conta corrente de Collor e da Ministra da Economia Zélia Cardoso de Melo.

Itamar Franco e sua equipe econômica acompanharam o período pré-Real que viria a ser implantado em 1994, sob a gestão do então Ministro da Fazenda e futuro governante Fernando Henrique Cardoso.

3.1.2 O Plano Real

O período de retomada do vigor produtivo no Estado pode ser observado no começo da década de 90, com uma ajuda por parte do cenário econômico, que veio por frear o processo inflacionário e permitir uma maior abertura comercial do país.

O Plano Real entrou em vigor a partir do dia 01/07/1994, sendo lançado por meio de uma Medida Provisória (MP) em 30/06/1994. O Plano Real segundo Sandroni (2002, p. 471), foi um dos planos que provocaram menores alterações para a economia, uma vez que seu lançamento foi precedido pela Unidade Real de Valor (URV) – moeda transitória entre o cruzeiro e o real –, com a finalidade de alinhar os preços e contribuir para um menor impacto dessa reforma monetária.

Ao contrário dos planos anteriores, o novo plano não veio acompanhado mais da banal medida de congelamento de preços, mas por medidas como: implante de uma nova moeda, o Real; fixa a taxa de câmbio na paridade de R\$ 1,00 para US\$ 1,00 – em meio ao Mercosul mais um problema com a apreciação do câmbio e o factível resultado de aumento de importação de produtos agrícolas dos membros do bloco –; acelera o processo das privatizações; eleva os juros²⁴; com a paridade se facilitam as importações; mantém a intenção de controle dos gastos públicos; dá continuidade no processo de abertura econômica e busca medidas de apoio à modernização das empresas além do estabelecimento de metas de inflação.

O Governo de Fernando Henrique tinha em pauta, além das diretrizes de uma menor intervenção deste no setor, tornar a atividade agrícola competitiva no mercado internacional, em termos de qualidade e preços de produtos. Segundo Brum (1998, p. 543), os produtores rurais deveriam se tornar empresários rurais eficientes, com informações e visão abrangente da cadeia produtiva e negócios, inseridos em um mercado altamente competitivo. O governo

²⁴ A âncora cambial teve segundo Alves & Paulillo (1998, p. 19), um papel danoso para a economia como um todo, mas particularmente perverso para a agricultura, pois eliminaria a possibilidade de inversões em produção e produtividade por parte dos produtores devido aos juros altos praticados.

ainda com problemas relacionados ao déficit público tomou, segundo Brum (1998, p.543), uma posição tímida e declinante, no critério subsídios para a agricultura.

Ainda Massuquetti (1998, p. 87) ratifica a idéia de Brum (1998) onde

[...] a economia passou por grandes transformações estruturais, o governo brasileiro tem procurado rever e redimensionar o papel do Estado em face às principais atividades econômicas. Nesse sentido, a agricultura é um dos setores mais atingidos, visto que, há alguns anos, o governo vem demonstrando clara intenção de diminuir a presença na atividade agrícola, deixando de ser o seu único financiador, maior regulador do mercado e controlador dos estoques físicos do produto, passando a estimular práticas mais modernas de financiamento e comercialização.

Porto Neto (1996) destaca duas medidas adotadas pelo governo: transição e o longo prazo. No que tange à transição, está a reformulação da Política de Garantia de Preços Mínimos, do crédito rural, dos estoques e do programa de seguro rural. Ainda, quanto à transição, cabe buscar a adaptação dos atuais instrumentos para uma realidade econômica de mercado e com menor intervenção estatal, onde, a soja demonstra ter já absorvido boa parte dessa mudança, pois existe uma boa interação da cultura com o mercado. Quanto às medidas de longo prazo, visavam uma redução do papel do Estado na regulação dos mercados. Entre as medidas de longo prazo, podem ser enquadradas

[...] reestruturação e fortalecimento do sistema de pesquisa e difusão de tecnologias, expansão, diversificação e modernização da infra-estrutura portuária e de transportes, reestruturação e fortalecimento do sistema de defesa agropecuária e a criação de mecanismos para facilitar e ampliar a participação do setor privado na comercialização (PORTO NETO, 1996, p. 10).

O Plano Real ainda teve uma última medida de peso, em 1999, promovendo a livre flutuação do câmbio e redução das taxas de juros domésticas. Fato que ao desvalorizar a moeda nacional em relação ao dólar, estimularia as exportações e a produção interna de máquinas e implementos agrícolas²⁵.

Fato indiscutível é que o Plano Real, em seu primeiro estágio, ao reduzir a inflação e neutralizar o componente de formação de preços, alterou o padrão de funcionamento estrutural do sistema econômico brasileiro. A mudança nos níveis de inflação ensejou o ressurgimento do potencial consumidor privado e ampliou o horizonte de planejamento para as decisões privadas.

O crescimento da agricultura brasileira, a partir de 1990, foi baseado, sobretudo, em ganhos consideráveis de produtividade, notadamente no setor de grãos, “a produtividade

²⁵ Do total de 146 fábricas existentes no Brasil, o que se pode observar é que o Estado de São Paulo detém 47,9% das mesmas, seguido do Rio Grande do Sul que detém 36,5%. O restante está distribuído entre os Estados do Paraná, Minas Gerais, Santa Catarina, Ceará, Rio de Janeiro e Alagoas (BNDES, 2003).

média do milho subiu 45%, do arroz 74%, do feijão 36%, da soja 23%, e do trigo 10%” (EMBRAPA, 2007). Sendo que a soja foi a única cultura a ter um crescimento expressivo na sua área cultivada ao longo das últimas três décadas puxado, pela demanda externa, o que evidencia uma expansão da fronteira agrícola, que atinge até a Amazônia atualmente (EMBRAPA, 2007).

No Plano Real vem ocorrendo um crescente avanço tecnológico na agricultura brasileira. Um indicador expressivo é o do consumo de fertilizantes, que evoluiu de 11,2 milhões de toneladas em 1994 para 14,3 milhões em 1999 (MF, 2004). Seguindo essa trajetória de desenvolvimento agrícola, o consumo de nutrientes e de matéria-prima também mostrou crescimento consistente desde a implantação do Real (IBGE, 2007).

Aliados ao plano vieram muitas outras políticas voltadas para a agricultura, para fins de uma recuperação do setor ainda muito importante, que responde por mais de um quarto das exportações nacionais.

O Governo Federal destinou recursos da ordem de R\$ 11,3 bilhões para o financiamento da safra 2000/2001, o que representa acréscimo de 36% em relação aos recursos de R\$ 8,3 bilhões programados para a safra anterior. Para apoiar o desenvolvimento de diversos setores da agricultura, o Plano Agrícola contempla novos programas com abrangência em todo o País, tais como os de fruticultura, agricultura e de apicultura. De forma complementar aos objetivos de incentivo à agricultura, melhorou as condições de captação de recursos por parte dos agricultores. Para o apoio à comercialização da safra, foram utilizados instrumentos como Contratos de Opção, Prêmio de Escoamento do Produto (PEP), Empréstimos e Aquisições do Governo Federal (EGF e AGF). As iniciativas governamentais promoveram níveis adequados de capitalização do setor rural, assim, foram decisivas para o aumento consistente da produção agrícola e do superávit comercial do setor. Esse superávit passou de US\$ 10,3 bilhões em 1992 para US\$ 13,4 bilhões em 1999.

Nos anos de 2001 e 2002, "os dispêndios na função Agricultura seriam menores que os observados nos últimos dezesseis anos” (GASQUES & VILLA VERDE, 2003, p. 152-153). E ainda, Gasques & Villa Verde (2003, p. 153) afirmam que essa redução de dispêndios ocorreria nas políticas de crédito rural e de sustentação de preços e de renda, justamente aquelas em que os países que protegem suas agriculturas mais concentram seus gastos, ajudando a desfavorecer o setor, quanto à concorrência externa e desleal aos países em desenvolvimento.

Os gastos dos estados seguiram o padrão da União, que teve dispêndios de aproximadamente 1,13% para a agricultura enquanto as médias dos estados ficaram em 1,20%

no ano de 2006. O Estado do Rio Grande do Sul ficou acima da média nacional investindo um total de 11 milhões de reais, ou seja, 2,07 % de suas despesas em agricultura, um percentual ainda muito baixo se comparado a Roraima 5,58% (MF, 2004). Gasques & Villa Verde (2003, p. 148), afirmam em seus estudos que o Estado deveria gastar no mínimo oito vezes mais do que gasta atualmente²⁶.

Para se ter uma idéia desse afastamento estatal, no período entre 1990 e 2001, o gasto público em agricultura em relação ao gasto total do Estado foi de 2,17%. E nos anos de 2000 e 2001 essa relação ficou ainda pior, por volta de 1%. Gasques & Villa Verde (1991) concluem que realmente existe uma redução entre os anos 80 e os últimos anos, ratificando uma mentalidade de progressivo desligamento do Governo para com a agricultura. Pode-se melhor observar essa involução nos gastos com auxílio do Tabela 3 que segue. Pois, fica claro que ao longo da década de 90 do século XX e início do XXI a redução de um ponto percentual das Despesas do Governo na Função Agricultura (DGFA) em relação às despesas totais da União. Ainda, se comparado ao PIB, a DGFA não tem acompanhado o crescimento do PIB.

Tabela 3 – Despesa global da união, PIB, DGFA e a participação do gasto em agricultura nos gastos totais 1990/2002 (*)

Anos	Despesa Total da União	Produto Interno Bruto	Despesa na Função Agricultura	Evolução do PIB (%) (**)	(C/A) %	(C/B) %
	(A)	(B)	(C)			
1990	709.848,60	898.407,40	13.769,80	100,00	1,53	1,94
1991	293.236,90	907.680,70	12.852,50	101,03	1,42	4,38
1992	368.603,10	902.742,30	9.249,10	100,48	1,02	2,51
1993	543.503,50	947.199,50	10.820,40	105,43	1,14	1,99
1994	513.704,30	1.002.637,90	11.399,30	111,60	1,14	2,22
1995	536.072,50	1.044.987,20	14.221,50	116,32	1,36	2,65
1996	453.998,5	1.072.769,20	9.880,30	119,41	0,92	2,18
1997	569.011,00	1.107.863,60	12.678,50	123,31	1,14	2,23
1998	699.115,60	1.109.325,10	9.384,90	123,48	0,85	1,34
1999	400.953,80	1.118.332,40	10.038,40	124,48	0,90	2,50
2000	680.246,20	1.167.116,80	6.815,50	129,91	0,58	1,00
2001	603.434,90	1.184.768,80	6.811,50	131,87	0,57	1,13
2002	301.714,00	1.346.027,60	6.068,90	149,82	0,45	2,01

Fonte: MF/STN e IPEADATA.

(*) Valores deflacionados pelo IGP-DI da FGV em milhões de reais.

(**) Base em 1990.

²⁶ Para Coelho (2002), apesar dos efeitos que esses dispêndios devem ter causado às contas públicas, esse valor ainda é uma cifra muito pequena se comparado ao que outros países gastam para proteger sua agricultura.

A agricultura, configurado um afastamento por parte do governo durante os anos 80 e na década de 90 passou a se relacionar de modo muito mais integrado, de um lado, com o sistema de distribuição, composto pela agroindústria e por cadeias de supermercados varejistas e de outro lado, com os fornecedores de insumos e produtos que estão cada vez mais ativos no mercado. No período em que realmente o crédito rural se torna escasso, essas novas fontes de financiamentos começam a ocupar este espaço – cadeias de supermercados, indústria de alimentos, *traders* e indústria de insumos. A economia cada vez mais tomou o lado das forças de mercado, cedendo espaço para a participação cada vez maior do capital privado no financiamento rural e mais atualmente, mudando a direção e as contas dos gastos públicos²⁷.

Gasques & Villa Verde (2003, p. 153) corroboram com essa afirmação

a principal razão da queda do gasto público em agricultura é a mudança introduzida pelo governo em relação aos instrumentos de política agrícola, como crédito rural e política de preços e estoques. No crédito rural, a participação do governo se dá através da equalização de taxa de juros. Na política de preços e estoques, as atribuições destas foram passadas em grande parte para a iniciativa privada.

Em 2002, o volume de crédito rural concedido – R\$ 9,7 bilhões de reais a preços de agosto de 1994 – foi equivalente a apenas 24,7% do que fora concedido em 1979, porém, apresentando uma melhora em relação ao ano anterior.

Outro fato que merece destaque fica por conta do papel do Banco Nacional Econômico e Social (BNDES). Tomando uma nova posição, com linhas de crédito para capital de giro e uma maior atenção para na reforma agrária, que pode ser visto como uma inovação, visto que é de tradição deste órgão o financiamento de equipamentos e máquinas. Então, segundo Alves & Paulillo (1998, p. 24) essa atual política de investimentos do governo distancia-se realmente dos empréstimos subsidiados e da concessão de facilidades fiscais para uma busca de novos instrumentos de financiamento, no mercado de capitais, por exemplo.

²⁷ Segundo Gasques & Villa Verde (2001), a análise dos gastos públicos pela ótica dos programas, permite verificar que ao longo destas últimas duas décadas, o governo distribuiu seus gastos em 30 programas, sendo que dois programas – abastecimento e administração – absorveram mais de 50% dos dispêndios. Entre os programas mais relevantes pela sua participação no gasto com abastecimento representou 41,30% dos dispêndios totais no período 1980-1999. Juntamente com administração, foram responsáveis por 57,1% dos dispêndios nesse período. A evolução dos gastos no programa de administração está ligada ao comportamento dos dispêndios do crédito rural e de outras políticas. O aumento da participação desse programa entre as décadas de 80 e 90, se deveu à transferência das contas da agricultura, referentes a fundos e programas que até 1987 estavam sob a responsabilidade do Banco Central, e que a partir de janeiro de 1988 passaram para a Secretaria do Tesouro Nacional. Deve-se chamar atenção para o fato de que esse aumento não representou uma expansão da máquina administrativa do governo, pois, em 1989, 82% dos dispêndios com administração pertenciam às operações oficiais de crédito.

As fontes de recursos do Sistema Nacional de Crédito Rural (SNCR) eram os saldos dos dispêndios à vista de bancos comerciais, os recursos dos fundos e programas agrícolas, os depósitos à vista do Banco do Brasil, entre outros. Como elucidado anteriormente, os recursos por parte do governo foram perdendo força, e aparecendo cada vez mais os financiamentos para custeio de safras. O Banco do Brasil em 1970 foi responsável por 79,19% na participação do crédito agrícola ofertado. Até meados dos anos 80, o Banco do Brasil ainda tinha um papel ativo na participação, mas, com as pressões no Governo foi se enquadrando nas novas políticas adotadas. O esgotamento do fluxo de poupança externa, da redução de poupança do setor público e da aceleração inflacionária foram os fatores que, segundo Massuquetti (1998), não permitiram a continuidade do modelo adotado anterior a década de 80.

Em decorrência disso, foram criadas novas fontes de recursos, com apoio crescente do setor privado. Dentre as novas fontes de financiamentos, surgidas pós-cruzado, podem ser citadas: o surgimento da Caderneta de Poupança Rural, Fundos Constitucionais, Depósitos Interfinanceiros Rurais (DIR), Fundos de Aplicações Financeiras (FAF), Sistema de Equivalência-Produto, Depósitos Especiais Remunerados (DER), os Recursos de origem externa e/ou de multinacionais, Mercado de Derivados, Certificado de Mercadoria com Emissão Garantida (CMG), Cédula do Produto Rural (CPR), Recursos Extramercado, Contratos de Compra e Venda de Soja Verde, Financiamento de Máquinas e Equipamentos (FINAME) dentre outras.

Após este sobrevôo sobre os últimos 60 anos de política agrícola e modernização da agricultura nacional, identifica-se que a agricultura brasileira sofrera alguns movimentos de mudanças conjunturais durante os últimos 20 anos. Conjunturais, relacionado ao comportamento da economia – inflação e abertura comercial – influenciando no crédito oferecido anteriormente. Cabe na subseção seguinte uma análise mais apurada e detalhada destes elementos de transição refletidos na produção e no Rio Grande do Sul.

3.2 Mudanças estruturais da agricultura

Nos anos 80, quando configurada uma crise fiscal do Estado, a estratégia de desenvolvimento agrícola baseado no crédito rural subsidiado se encontrava em decadência, alterando o modelo que até então vinha sendo defendido. Com uma revisão e um maior controle dos gastos públicos, as tendências dos dispêndios relacionados com a agricultura

apontavam para uma redução ascendente, o que acaba por diminuir os subsídios largamente disponibilizados.

O setor agrícola passou então, a experimentar o gradual afastamento do governo e a viver os dismantelamentos dos instrumentos de intervenção na agricultura que foram intensamente utilizados desde os anos 60 e 70.

A partir de 1983, com o desaparecimento do principal mecanismo de compensação da agricultura – subsídio por parte do crédito rural – o crescimento da produção e o aumento da capacidade competitiva brasileira, que sempre estiveram associados aos avanços de fronteira agrícola, migram para uma nova característica que não mais seriam os aumentos horizontais.

A agricultura então, se ajustando às restrições que enfrentou em 80 e 90, buscava formas alternativas de capitalização interna ao estabelecimento produtivo, ligados principalmente ao forte aumento da produtividade (DIAS & AMARAL, 2001). Rangel (2004, p. 40) decreta que “gradativamente a agricultura se converte na atividade que é nos países mais desenvolvidos”, ou seja, fornecimento de matéria-prima para uma série de indústrias específicas.

Mas, em detrimento de fatores elucidados anteriormente, se faz necessária uma breve análise da evolução da produção e produtividade gaúcha e os novos métodos utilizados para driblar as dificuldades com o distanciamento do Estado da agricultura.

3.2.1 A Produção e Produtividade no Rio Grande do Sul

À medida que se desenvolvia a qualidade dos produtos, graças à aplicação de novos procedimentos e novas técnicas de produção, nascia o sentido de eficácia e eficiência (ROMANO, 1987).

A estabilidade nas áreas plantadas, verificadas nas duas últimas décadas junto aos aumentos do volume de produção, aponta para uma maior preocupação por parte do produtor rural em investir nos meios de tecnologia que proporcionem aumento de rendimento, do que no aumento de fronteira, característica que predominou até meados dos anos 70, como já mencionado.

Até a década de 50, as culturas mais desenvolvidas, como arroz e trigo – plantado desde o século XIX – sofriam sérios problemas de competitividade externa e queda nos

preços. Tal situação favoreceu a inserção da sojicultura no Estado. Schneider & Fialho (2000, p. 127) apontam que a frente da expansão da soja localizou-se na zona colonial do noroeste do Estado, incluindo Missões, Alto Uruguai e Planalto Médio, tendo sido sua difusão baseada nos “incrementos anuais de área plantada”. Período em que os recursos disponíveis pelo governo eram abundantes e revertidos em incorporações de novas fronteiras de produção com investimentos em tecnologia e “insumos modernos” que possibilitassem tais aumentos.

Considerando “que o tamanho do solo é uma das muitas atividades que absorvem o tempo e os recursos do agricultor” (RANGEL, 2004, p. 40), então, frente às dificuldades encontradas no período estudado, o produtor se viu obrigado a racionalizar o uso das áreas agrícolas, preocupando-se mais com a produtividade da terra e redução dos custos do que com a conquista de novas áreas para cultivo. Essa etapa de preocupação voltada para o rendimento, e mais efetiva após os anos 90, marcou segundo Teixeira (2004) uma nova etapa do processo histórico de modernização da agricultura, posterior aos complexos agroindustriais.

As transformações ocorridas no noroeste, fizeram com que o modo de vida colonial fosse substituído pela cultura tecnificada e integrada aos complexos agroindustriais (SCHNEIDER & FIALHO, 2000, p. 128). Nas regiões sul e oeste do Rio Grande do Sul onde predominava a produção de arroz irrigado e a pecuária extensiva, a dinâmica foi diferente. Ao passo que, na década de 80, a produção de arroz acompanhou o processo de aumento de produtividade, a pecuária se manteve estagnada (SCHNEIDER & FIALHO, 2000, p. 128).

Schneider & Fialho (2000, p. 127) salientam que, nestas regiões onde se desenvolveu a soja, a modernização tecnológica se revelou intensa e, em razão disso, um grande número de colonos viu-se forçado a abandonar suas atividades rurais, procurando alternativas. Ainda, Schneider & Fialho (2000, p. 131) apontam que a categoria dos empregados agrícolas vem diminuindo no Rio Grande do Sul, assim como, os agricultores familiares, também continuam a reduzir sua presença no meio rural gaúcho.

Mas, não é correto pensar em uma falta de espaço para essa atividade de menor capitalização, tão pouco, significa dizer que modernização e agricultura familiar são termos excludentes mesmo que muito se discuta sobre isto²⁸. Pois, conforme Prieb (2005, p. 117) “o rural não pode mais ser visto como eminentemente agrícola”. Pois, na Região do Vale do Rio

²⁸ Para uma leitura sobre o assunto existem diversas literatura, uma boa opção seria começar por Alberto Passos Guimarães, Ignácio Rangel, Costa Delgado e José Graziano da Silva.

Pardo onde predomina a fumicultura – de caráter familiar – existe uma forte articulação entre produtor e indústria²⁹.

Tal sucesso na relação se reflete no fato de que o Brasil entre 1999-2002 figurou como maior exportador de fumo em folha. Prieb (2005, p. 125) mostrou em sua pesquisa que 43% de sua amostra declaram que a Renda familiar era 50% ou mais de origem agrícola. Fechando essa idéia, Lopes & Rocha (2005, p. 32) em estudo realizado com os dados censitários do Censo Agropecuário 1995/96 do Instituto Brasileiro de Geografia e Estatística (IBGE) salientam que a contribuição da agricultura familiar mostra-se importante no valor de produção de fumo (86,3%), mandioca (73,2%) e horticultura (42%).

Dessa forma, fica evidente o fato de que, apesar do Brasil não ter subsídios explícitos, a produtividade cresce. Por exemplo, a soja brasileira cresce mais do que a dos norte-americanos, contribuindo assim para rebaixamento nos custos de produção³⁰ (EMATER, 2007).

A abertura econômica que se iniciou a partir do Plano Collor, marca o aumento gradual nas perspectivas de globalização. Esse fator traz novas opções de insumos e conseqüentemente uma maior competição do setor em relação às pressões externas.

A abertura comercial fez com que o produtor brasileiro, ajustasse seus fatores de produção, seus custos bem como entendesse a importância da formação da receita, que em última análise definem a renda líquida do produtor. Internamente, o Brasil precisa avançar na melhoria do custo Brasil e outros fatores que oneram a cadeia produtiva da soja, fazendo com que essa competitividade conquistada pela eficiência na produção de soja, não seja maior pela debilidade das políticas públicas no apoio ao setor (EMATER, 2007).

Mesmo com o acirramento da concorrência, estes fatores vieram colaborar com o processo de aumento da produção, que para continuar no mercado, os métodos de produção nacional deveriam se moldar ao modelo adequado-racional e se tornar mais competitivos.

Teixeira (2004) afirma que a agricultura nacional nunca esteve tão próxima do modelo dos países desenvolvidos, no quesito produção/produtividade. Conforme Teixeira (2004), nos

²⁹ Prieb (2005, p. 39) salienta que diferentemente do que ocorre na maior parte dos países produtores de fumo, no Brasil, esse processo agrega dois componentes que são essenciais para o sistema de integração – a agroindústria e a agricultura familiar – o que denota especificidades próprias.

³⁰ No Rio Grande do Sul, o custo da soja para safra 2002/03, onde o custo por saca baseado no dólar em primeiro de julho no valor de US\$ 2,859, o custo total por saca ficou em US\$ 6,13 a saca, contra US\$ 5,77 na safra passada. Os fatores que fizeram com que a soja tenha alcançado boa produtividade foram: tecnologia, investimento em pesquisa, ajustes do parque de máquinas, racionalização da gestão de custos e também fatores climáticos. Os produtores esperam que mudanças ocorram principalmente, nas políticas públicas que possam contribuir para que o país no futuro se torne insuperável na competitividade da soja no mercado mundial (EMATER, 2007).

países de primeiro mundo, a tecnologia vem como fator mais importante na produção, ou seja, a produção deve crescer com o aumento sustentado da produtividade do trabalho na terra.

Segundo Zambonadi (1996, p. 29), dados oficiais revelam uma verdadeira revolução em termos de ganho na produtividade na agricultura brasileira nos últimos dez anos, especialmente se forem considerados a redução gradual dos volumes de recursos aplicados em crédito rural e pesquisa no setor, variáveis que inibiram o investimento contínuo em novos equipamentos e de incorporações de novas áreas para produção.

Uma das variáveis mais importantes que auxiliaram no crescimento da produtividade pode ser atribuída à disponibilidade do estoque e ao domínio de tecnologias geradas pela pesquisa e assistência técnica nas décadas de 70 e 80. Importante foi, sem dúvida, o papel da EMBRAPA, com o aproveitamento por parte dos produtores de tecnologias desenvolvidas pela empresa estatal responsável pela pesquisa e avanços agrícolas/tecnológicos. Sua função, também, se estende a outras funções, exercendo papel também de fomentador da demanda interna pela soja, com programas como Soja na Mesa e Soja Dez, Fome Zero, cujo principal objetivo era o de divulgar os benefícios para a saúde e alimentação e seu teor altamente protéico.

O agricultor aumentou a produtividade da lavoura e ajustou-se a uma nova realidade, com maior risco econômico e de rentabilidade duvidosa, em face à queda dos preços³¹ e da instabilidade das políticas internas de sustentação de renda, que foram eficientes até a década de 80 e pouco operantes nas safras mais recentes. A redução da lucratividade das lavouras obrigou o produtor a operar mais com o espírito empresarial do que com a emotividade que lhe era peculiar.

Explicar o aumento de produtividade na agricultura, com a queda de renda, e a capacidade de poupar do setor, pode estar ligado a outros fatores, também, tais como a redução do custo global da produção e a inclusão de algumas tecnologias específicas, com uso de insumos mais modernos, que além do bom comportamento climático possam justificar ganhos tecnológicos e aumento de produtividade.

Segundo Teixeira (2004) a inserção do país no bojo do projeto neoliberal, trouxe consigo mudanças para o setor de sementes e agroquímicos, ocorrendo um intenso fenômeno transfronteiras de fusões, aquisições e incorporações pelo capital internacional. Concomitantemente a essa realidade o Brasil ingressa na era da biotecnologia, sendo motivo

³¹ Müller (1999) salienta esta observação, acrescentando que nos anos 80 os preços recebidos pelos produtores de soja mostraram-se tendencialmente decrescentes, com destaque maior para o período posterior a 1985. Onde esse fato se explique que devido ao aumento de produtividade da terra levariam a uma diminuição no preço dos produtos agrícolas.

de discussões acaloradas, nos últimos anos no país e no Estado, especialmente. Pretende-se então, na seção seguinte, abordar revisão de literatura da metodologia que permitiu a realização do presente trabalho e concretização dos objetivos. Pois, a análise multivariada possibilitou agrupar padrões de níveis de modernização da agricultura gaúcha e ainda analisar e hierarquizar estes diferentes grupos por um número considerável de variáveis exploratórias.

4 ANÁLISE ESTATÍSTICA MULTIVARIADA DE DADOS

Hair Jr. (2005) atribui a Wilks³² a formulação do construto teórico da análise multivariada de variância. Contudo, lembra que só foi possível sua utilização de forma mais prática pelos pesquisadores com o desenvolvimento dos computadores e do aperfeiçoamento dos *softwares*. Kendall (1957, apud REIS, 1997, p. 21) define análise multivariada como sendo um ramo da análise estatística o qual verifica a relação entre grupos de variáveis dependentes e “coletadas sobre o mesmo indivíduo” (MORRISON, 1976, p. 22).

As n observações sobre p variáveis constituem a chamada matriz de dados, em que, $n > p$. Assim, a exigência para aplicação é que o número de variáveis seja menor que o de observações e, ainda, recomenda-se “a análise de um mínimo [...] de 20 observações” (HAIR Jr., 2005, p. 282).

É fato considerar que “a classificação é uma atividade conceitual básica dos seres humanos” (REIS, 1997, p. 287). Entretanto, as Ciências Sociais e Biológicas, a Medicina ou Arqueologia se utilizam de critérios de classificação na tentativa de explicar a complexidade com que se depara o pesquisador em meio a um elevado número de observações e variáveis. Então, as facilidades oferecidas pelo desenvolvimento dos meios computacionais e dos pacotes estatísticos têm possibilitado o uso crescente das técnicas de análise estatística como suplemento e instrumento de enorme utilidade para os pesquisadores. Assim, o presente trabalho pretende explorar a técnica conhecida como Análise Multivariada de dados.

Para espanto de muitos, o artigo seminal utilizando tal metodologia teve origem na Psicologia. Pois, foi criada/utilizada para classificar/encontrar diferentes padrões de personalidade nos pacientes com base em questionários (HOFFMANN, 1999, p. 3). O caso é que pela funcionalidade em otimizar a interpretação de grandes conjuntos de dados, pretende-se adotar este método na análise econômica, em especial, para verificar os níveis de modernização da agricultura nos COREDEs do Rio Grande do Sul (RS) tendo como fonte de dados o Censo Agropecuário de 1995/96 e, assim, difundir essa importante/funcional ferramenta para utilização na área das Ciências Econômicas.

A análise multivariada, de regra, na literatura está dividida em análise de agrupamentos, análise fatorial das componentes principais, dentre outros. Estando de acordo com o sentido do método, este trabalho segue ordinalmente a apresentação sumária supra-exposta de cada

³² WILKS, S.S. Certain Generalizations in the Analysis of Variance. *Biometrika*, 1935, n.24, p.471-94.

técnica, a começar, pelo agrupamento dos *cases* propostos para o trabalho, passando a análise fatorial das respectivas componentes principais.

4.1 Análise de Agrupamento (AA)

A Análise de Agrupamento ou *Cluster Analysis* é um método que tem como objetivo, como o nome já denota, agrupar indivíduos – *cases* – ou variáveis em grupos discretos. “A análise é realizada com base na similaridade ou dissimilaridade entre as variáveis a serem analisadas” (JOHNSON & WICHERN, 1992, p. 573). Sendo assim, estes se agruparão por apresentarem características similares de acordo com algum critério de classificação. Ainda, Härdle & Simar (2003, p. 301) definem AA como um conjunto de ferramentas para construir grupos – *clusters* – de um banco de dados multivariado de observações e variáveis.

Naturalmente, doravante o agrupamento por semelhança obtém-se grupos com elementos homogêneos entre si, enquanto os demais grupos devem ser, o mais próximo possível, desiguais (VALENTIN, 2000). Sendo assim, a diferença entre os vários grupos deve ser o mais distante possível. Para tanto, “quanto mais próximo estão os elementos de um grupo, mais homogêneos eles serão”, arrematam (HÄRDLE & SIMAR, 2003, p. 301).

Finalmente, o pesquisador se depara com diversos métodos de aglomeração, assim como, medidas de distância entre os grupos a serem formados. Portanto, discute-se primeiramente acerca das medidas de distâncias encontradas na literatura usual de técnicas multivariadas. Na subsecção seguinte, apresentam-se as técnicas de AA, bem como, a escolha do método e coeficiente de medida a serem utilizados pelo trabalho proposto.

4.1.1 As medidas de similaridade ou dissimilaridade

Para aplicação da AA é necessário a escolha de um coeficiente que quantifique o quão parecidos dois ou mais indivíduos são. Sendo que este coeficiente diz respeito à estimação de uma medida de similaridade/dissimilaridade entre estes indivíduos e/ou populações a serem agrupados. Na medida de similaridade, quanto maior o valor observado, mais parecidos serão

os indivíduos. Enquanto na dissimilaridade, o contrário se verifica, ou seja, quanto maior o valor observado, menos parecidos estes serão.

Assim sendo, pode-se apontar o coeficiente de correlação como uma medida de similaridade, pois quanto maior seu valor, maior a associação entre os indivíduos e, de dissimilaridade a distância Euclidiana, que quanto menor o valor, mais próximo os objetos estarão e homogêneos serão. Para que seja possível a escolha do melhor coeficiente de semelhança, é necessário ter uma matriz de dados $X_{(n \times p)} = X_{ij}$. Assim, cada vetor linha indica um *case* (indivíduos, espécies, tratamentos, etc), e cada vetor coluna, uma variável (REGAZZI, 2001), o que fica melhor vislumbrado com o auxílio da Tabela 4.

Tabela 4 – Matriz de dados n indivíduos e p variáveis.

Indivíduos	Variáveis					
	X ₁	X ₂	X ₃	X ₄	...	X _p
1	X ₁₁	X ₁₂	X ₁₃	X ₁₄	...	X _{1p}
2	X ₂₁	X ₂₂	X ₂₃	X ₂₄	...	X _{2p}
3	X ₃₁	X ₃₂	X ₃₃	X ₃₄	...	X _{3p}
⋮	⋮	⋮	⋮	⋮	...	⋮
n	X _{n1}	X _{n2}	X _{n3}	X _{n4}	...	X _{np}

Fonte: Elaboração própria.

O primeiro passo para que se inicie uma AA é a conversão da matriz $n \times p$ de dados em uma matriz quadrada, onde n é o número de indivíduos e p o número de variáveis observadas. Dado o valor de um conjunto de p variáveis, em cada intersecção da i -ésima linha, e da k -ésima coluna dessa matriz, coloca-se a medida de similaridade/dissimilaridade, entre o i -ésimo e k -ésimo indivíduo. A alta similaridade indica que dois indivíduos são comuns em relação ao conjunto de variáveis, enquanto que a alta dissimilaridade indica o contrário.

Sendo o objetivo da AA reunir objetos o mais semelhantes possível, torna-se necessário eleger alguma medida para avaliar o quão semelhantes, ou diferentes são os indivíduos. Vicini (2005, p. 13) sugere “avaliar a semelhança em termos de distância entre pares de objetos”.

A distância Euclidiana é, com toda certeza, a mais difundida e utilizada na análise de agrupamentos. Contudo, vale lembrar, que é recomendável padronizar as variáveis antes de

calcular esta distância. Pois, a distância Euclidiana preserva a distância relativa quando as variáveis estão padronizadas.

Esta é utilizada para calcular medidas específicas, assim como a distância Euclidiana simples e a distância Euclidiana média. Para formas de agrupamento Ward e Centróide, a distância euclidiana quadrática é a mais recomendada (HÄRDLE & SIMAR, 2003, p. 308). Então, seguindo o raciocínio supra-exposto, quanto menor a distância Euclidiana dos indivíduos, mais semelhantes serão. Esta é obtida mediante o teorema de Pitágoras, para um espaço multidimensional, podendo ser definida como:

$$d_{AB} = \sqrt{(X_{1A} - X_{1B})^2 + (X_{2A} - X_{2B})^2 + \dots + (X_{jA} - X_{jB})^2} \quad (1.1)$$

onde A e B são os indivíduos e j as variáveis.

A distância Euclidiana média é derivada da expressão anterior e também bastante utilizada. Possui duas considerações importantes: a) ela pode ser utilizada na ausência de dados para algumas coordenadas e b) permite acumular evidências empíricas sobre níveis de similaridade. A distância entre dois agrupamentos é obtida pela média das distâncias. Ou seja:

$$dem_{AB} = \sqrt{\frac{1}{j} [(X_{1A} - X_{1B})^2 + (X_{2A} - X_{2B})^2 + \dots + (X_{jA} - X_{jB})^2]} \quad (1.2)$$

A distância Euclidiana quadrática é expressa por:

$$deq_{AB}^2 = (X_{1A} - X_{1B})^2 + (X_{2A} - X_{2B})^2 + \dots + (X_{jA} - X_{jB})^2 \quad (1.3)$$

Além destas três distâncias, a literatura reúne mais as distâncias de Mahalanobis, Minkowski, Manhattan, Chebychev, Potência e o Coeficiente de Pearson ou Correlação de Pearson. Para tanto, o presente trabalho fará uso da distância Euclidiana quadrática e, em vista disso, limita-se apenas a elencar as outras distâncias que poderão ser citadas em algum tipo de trabalho de AA. Sendo assim, para outros tipos de distâncias, sugere-se os trabalhos de Mardia, Kent & Bibby (1979), Manly (1986), Johnson & Wichern (1992) e Vicini (2005) que expõem tais distâncias. Por fim, escolhida a medida de dissimilaridade, o passo seguinte se restringe a mais uma escolha, a do método para a Análise de Agrupamentos.

4.1.2 Os métodos de Análise de Agrupamento

Assim como na escolha entre as diversas medidas de dissimilaridade e similaridade, o pesquisador se depara igualmente em semelhante posição, ou seja, da escolha do método adequado para se realizar uma AA.

É possível dividir as técnicas disponíveis em vários grupos, no entanto, derivam de dois grandes denominados de Hierárquicos e Não-Hierárquicos. Do ponto de vista hierárquico, têm-se: a) Divisivos e b) Aglomerativos. Já para o processo que dispensa a condição hierárquica, dividem-se em: a) Limiar Sequencial; b) Limiar Paralelo e c) Particionamento Otimizador (MALHOTRA, 2001, p. 531). Contudo, o presente trabalho se detém a abordar apenas os métodos hierárquicos, pois, na execução do mesmo será utilizada a forma aglomerativa.

A técnica hierárquica caracteriza-se pelo estabelecimento de uma hierarquia que usualmente é apresentada na forma de uma árvore denominada dendrograma. No aglomeramento divisivo, parte-se de um grande grupo que contém todos os indivíduos/objetos da pesquisa, que através de um processo sistemático de divisões sucessivas, obtêm-se n grupos de um elemento cada.

Já o “método mais difundido e comumente utilizado é a forma aglomerativa” (REIS, 1997, p. 298) e, por este motivo, será o utilizado para alcançar os objetivos propostos por este estudo. Na forma aglomerativa, faz-se o inverso, isto é, parte-se de n grupos de apenas um indivíduo cada, que vão sendo agrupados sucessivamente até que se encontre um grupo que incluirá a totalidade dos n indivíduos ou variáveis em questão.

Ainda, os procedimentos hierárquicos usam cinco algoritmos aglomerativos diferentes para desenvolver os agrupamentos, os métodos de ligação individual, ligação completa, ligação média, o método de Ward e o método centróide.

4.1.2.1 O Método da Ligação Simples

O método de ligação individual ou simples – ainda pode ser encontrado como o método do vizinho mais próximo e *Single Linkage*³³ – é baseado na distância mínima. É o

³³ O presente trabalho limita-se a descrever com mais afinco e detalhamento o *Single Linkage* e o Método de Ward. Pela justificativa de que, o primeiro, é um dos algoritmos mais conhecidos/utilizados e, o segundo, por

método no qual os grupos são, inicialmente, constituídos cada um de um indivíduo e reunidos de acordo com a proximidade dos elementos, sendo os indivíduos mais próximos fundidos sucessivamente até formar-se um único grupo. Então, a idéia de distância mínima, na qual o método se baseia, pode ser vislumbrada com a Figura 1:

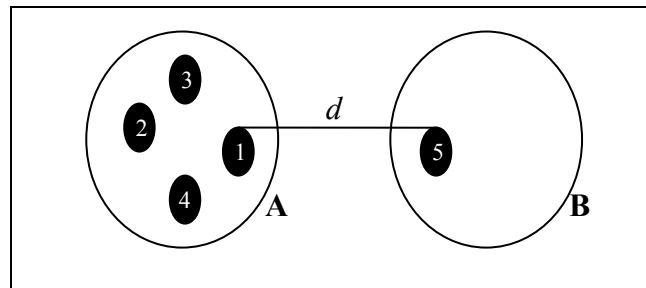


Figura 1 – Distância mínima entre grupos.

Doravante a figura, surge a necessidade de se ter escolhido a distância³⁴ a ser utilizada que, assim, irá gerar uma matriz – seguindo os indivíduos/variáveis da Figura 1 ter-se-ia uma matriz $D_{5 \times 5}$ – de distâncias entre todos os cinco indivíduos/variáveis. Para que seja possível formar grupos com características homogêneas entre os indivíduos da figura anterior, foram selecionadas quatro variáveis com valores hipotéticos, a fim de auxiliar com um exemplo a forma de agrupamento pelo Método Hierárquico Aglomerativo, através do algoritmo do *vizinho mais próximo* pela distância Euclidiana. Assim, para ilustrar o método, tomam-se os valores das observações acerca dos cinco indivíduos que estão sumariamente expostos na Tabela 5.

Tabela 5 – Número de indivíduos e suas respectivas variáveis.

Indivíduo	Variável 1	Variável 2	Variável 3	Variável 4
1	20	18	11	10
2	5	9	35	3
3	11	10	30	7
4	7	2	15	4
5	49	45	7	26

Fonte: Elaboração própria.

que será utilizado no presente estudo, contudo, os outros métodos serão, apenas, sucintamente descritos a fim de introduzi-los ao leitor.

³⁴ A escolha e conceitos de distâncias ou dissimilaridades da subseção anterior.

Para saber quais são as menores distâncias entre os indivíduos e dar início a formação dos grupos, calcula-se essas através da equação 1.1:

$$d_{ind1,ind1} = \sqrt{(20-20)^2 + (18-18)^2 + (11-11)^2 + (10-10)^2} = 0;$$

$$d_{ind1,ind2} = \sqrt{(5-20)^2 + (9-18)^2 + (35-11)^2 + (3-10)^2} = 30,5;$$

$$d_{ind1,ind3} = \sqrt{(11-20)^2 + (10-18)^2 + (30-11)^2 + (7-10)^2} = 22,7;$$

$$d_{ind1,ind4} = \sqrt{(7-20)^2 + (2-18)^2 + (15-11)^2 + (4-10)^2} = 21,8;$$

$$d_{ind1,ind5} = \sqrt{(49-20)^2 + (45-18)^2 + (7-11)^2 + (26-10)^2} = 42,9;$$

$$d_{ind2,ind3} = \sqrt{(11-5)^2 + (10-9)^2 + (30-35)^2 + (7-3)^2} = 8,8;$$

$$d_{ind2,ind4} = \sqrt{(7-5)^2 + (2-9)^2 + (15-35)^2 + (4-3)^2} = 21,3;$$

$$d_{ind2,ind5} = \sqrt{(49-5)^2 + (45-9)^2 + (7-35)^2 + (26-3)^2} = 67,4;$$

$$d_{ind3,ind4} = \sqrt{(7-11)^2 + (2-10)^2 + (15-30)^2 + (4-7)^2} = 17,7;$$

$$d_{ind3,ind5} = \sqrt{(49-11)^2 + (45-10)^2 + (7-30)^2 + (26-7)^2} = 59,7$$

$$d_{ind4,ind5} = \sqrt{(49-7)^2 + (45-2)^2 + (7-15)^2 + (26-4)^2} = 64,5.$$

Tendo sido calculadas todas as distâncias, obteve-se a seguinte matriz quadrada de distâncias Euclidianas entre os indivíduos:

$$D_1 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 30,5 & 22,7 & 21,8 & 42,9 \\ \dots & 0 & \mathbf{8,8} & 21,3 & 67,4 \\ \dots & \dots & 0 & 17,7 & 59,7 \\ \dots & \dots & \dots & 0 & 64,5 \\ \dots & \dots & \dots & \dots & 0 \end{bmatrix} \end{matrix}$$

Assim, os dois primeiros indivíduos, os mais próximos – com menor distância –, são fundidos e formam o grupo I que contém os indivíduos 2 e 3 com distâncias $d_{23} = 8,8$.

Seguindo o raciocínio, tem-se uma nova matriz, D_2 :

$$D_2 = \begin{matrix} & 1 & \mathbf{(23)} & 4 & 5 \\ \begin{matrix} 1 \\ \mathbf{(23)} \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 22,7 & 21,8 & 42,9 \\ \dots & 0 & \mathbf{17,7} & 59,7 \\ \dots & \dots & 0 & 64,5 \\ \dots & \dots & \dots & 0 \end{bmatrix} \end{matrix}$$

Busca-se, agora, o vizinho mais próximo ao grupo I, aquele de menor distância dos grupos individuais 1,4 e 5 com base na matriz D_2 :

$$d_{(23)1} = \min \{d_{21}; d_{31}\} = \min \{30,5; 22,7\} = \min d_{13} = 22,7;$$

$$d_{(23)4} = \min \{d_{24}; d_{34}\} = \min \{21,3; 17,7\} = \min d_{43} = 17,7;$$

$$d_{(23)5} = \min \{d_{25}; d_{35}\} = \min \{67,4; 59,7\} = \min d_{53} = 59,7.$$

Observa-se que $d_{(23)4} < d_{(23)1} < d_{(23)5}$, logo, a segunda menor distância está na linha 23 e coluna 4, representada em D_2 por $d_{(23)4} = \min d_{43} = 17,7$. Então, o indivíduo 4 será incluído ao grupo já existente, formado por 2 e 3, o que dá origem ao grupo II. Novamente, dado o segundo passo na aglomeração, se reconstrói a matriz, agora D_3 , após o novo passo do agrupamento:

$$D_3 = \begin{matrix} & 1 & \mathbf{(234)} & 5 \\ \begin{matrix} \mathbf{1} \\ \mathbf{(234)} \\ 5 \end{matrix} & \begin{bmatrix} 0 & \mathbf{21,8} & 42,9 \\ \dots & 0 & 59,7 \\ \dots & \dots & 0 \end{bmatrix} \end{matrix}$$

Na busca do vizinho mais próximo ao grupo II, tendo por base D_3 , de forma análoga a anterior:

$$d_{(234)1} = \min \{d_{(23)1}; d_{41}\} = \min \{22,7; 21,8\} = \min d_{14} = 21,8;$$

$$d_{(234)5} = \min \{d_{(23)5}; d_{45}\} = \min \{59,7; 64,5\} = \min d_{(23)5} = 59,7.$$

Como $d_{(234)1} < d_{(234)5}$, a menor distância está na linha 1 e coluna 234, o que equivale a dizer que $d_{(234)1} = \min d_{41} = 21,8$. Assim, o grupo III agora formado pelos indivíduos 1,2,3 e 4 e, a matriz D_4 é obtida de forma análoga às anteriores:

$$D_4 = \begin{matrix} & (1234) & 5 \\ (1234) & \begin{bmatrix} 0 & 42,9 \\ \dots & 0 \end{bmatrix} \\ 5 & & \end{matrix}$$

Finalmente, agrega-se o último indivíduo, o mais distante – $d_{(1234)5} = \min \{d_{(234)5}, d_{15}\} = \min \{59,7, 42,9\} = \min d_{15} = 42,9$ para formar o grande grupo que deverá conter todas as variáveis.

Conforme Valentin (2000, p. 56), para formar e interpretar o dendrograma segue-se a lógica dos itens:

- no eixo vertical são dispostos os valores das distâncias Euclidianas;
- para compor o dendrograma, busca-se na matriz D de distâncias o menor valor, ou a menor distância, isto é, uma menor dissimilaridade entre os indivíduos. Sendo assim, a menor distância é a de 8,8. Está entre as variáveis 2 e 3, que foram reunidas no dendrograma na altura 8,8 do eixo vertical formando, assim, o primeiro grupo;
- seguindo a lógica do *Single Linkage*, a segunda menor distância entre os grupos individuais e o grupo I anteriormente formado, é o indivíduo 4. Este distante 17,7 e agora reunido ao primeiro grupo, formando, assim, o grupo II;
- o indivíduo 1 agora é o que possui a menor distância do grupo formado por 2, 3 e 4. Sua distância é de 21,8 e, se agrupa às variáveis do grupo I e do grupo II, formando o grupo III;
- por fim, o indivíduo mais distante, isto é, 42,9 do indivíduo 1, que faz parte do grupo III, este se agrupa a 1, 2, 3 e 4 para formar o último grupo, que contém todos os indivíduos do exemplo, dando origem ao dendrograma completo ilustrado na Figura 2.

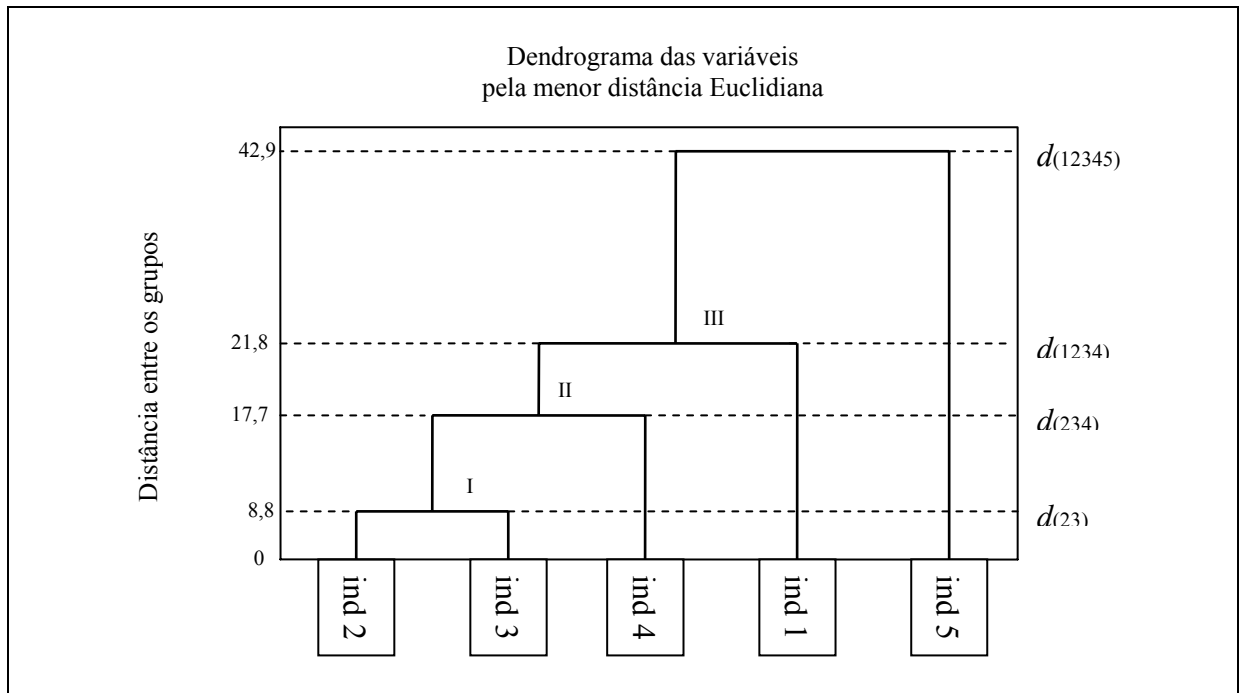


Figura 2 – Dendrograma referente ao agrupamento pelo método da ligação simples através da distância Euclidiana.

Em razão de sua simplicidade, o método *Single Linkage* apresenta uma desvantagem. O fato de reunir um objeto ao elemento mais próximo do grupo já formado, faz com que os objetos intermediários entre os grupos sejam rapidamente aglomerados a estes tornando os grupos menos homogêneos. Ocorre um encadeamento de objetos que dificulta a separação dos grupos e tende a manter isolados outros elementos que ainda não foram anexados. Assim, Valentin (2000) sugere que, para estudos em que as amostras de características intermediárias sejam numerosas, este método seja evitado.

4.1.2.2 O Método da Ligação Completa

Além do vizinho mais próximo, existe o método da ligação completa – *Complete Linkage* – ou método do vizinho mais distante, que possui um procedimento inverso ao anterior.

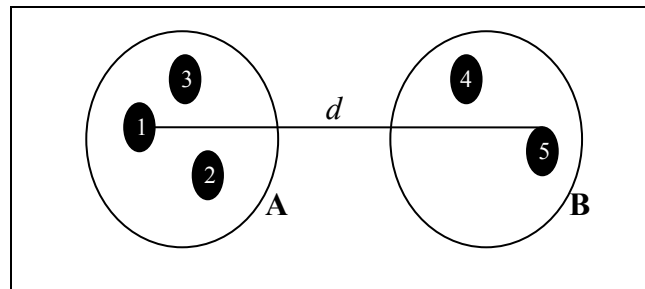


Figura 3 – Distância máxima entre grupos.

A distância entre dois grupos é agora definida como sendo a distância entre os elementos mais afastados/distantes ou menos semelhantes. Dados dois grupos de elementos (i, j) e (k) , a distância entre eles é a maior das distâncias entre seus elementos, ou seja, $d_{(i,j)k} = \max\{d_{ik}; d_{jk}\}$. Assim, cada grupo passa a ser definido como um conjunto de elementos em que cada um é mais semelhante a todos os restantes elementos do grupo do que a qualquer dos elementos dos grupos restantes.

Como desvantagem, o *Complete Linkage*, tende a formar *clusters* compactos compostos de indivíduos muito semelhantes entre si. Embora os resultados da aplicação deste método dêem uma visão muito nítida dos grupos formados, nem sempre apresentam um elevado grau de concordância com a estrutura inicial dos dados.

4.1.2.3 O Método das Médias das Distâncias

É encontrado, também, como *Average Linkage* na literatura. Utiliza a média das distâncias entre todos os pares (cada elemento de um par pertence a grupos distintos) de elementos formados a partir dos dois grupos, para gerar a nova distância a ser utilizada na matriz de distâncias. Ou seja, um grupo passa a ser definido como um conjunto de indivíduos no qual cada um tem mais semelhanças, em média, com todos os membros do mesmo grupo do que com todos os elementos de qualquer outro grupo.

Reis (1997, p. 317) salienta que este método forma grupos mais homogêneos do que os formados pela Ligação Simples, contudo, menos homogêneos que os formados através da Ligação Completa. Pois, enquanto que no “vizinho mais próximo” e no “vizinho mais distante” a inclusão de um novo indivíduo num grupo depende de um único valor de

semelhança, o menor ou o maior, respectivamente, a estratégia da média, por tomar valores médios, tem a vantagem de evitar valores extremos e de tomar em consideração toda a informação dos grupos.

4.1.2.4 O Método do Centróide

Nesta opção de algoritmo de agrupamento, a distância entre dois grupos é definida como a distância entre os seus centróides³⁵, pontos definidos pelas médias das variáveis caracterizadoras dos indivíduos de cada grupo, isto é, o método calcula a distância entre dois grupos como a diferença entre as médias, para todas as variáveis.

A cada agrupamento de novo indivíduo, deve-se calcular o novo centróide. E, por este motivo, gera certa dificuldade, pois, a todo o momento os dados originais devem ser resgatados. Este método e o das Médias das Distâncias são muito parecidos e, por conta disso, na maioria das vezes apresentam agrupamentos semelhantes.

4.1.2.5 O Método de Ward³⁶

Caracteriza-se pela “formação de grupos com alta homogeneidade interna e muito apropriado para variáveis quantitativas” (MINGOTI, 2005, p. 178). Também chamado de Mínima Variância, este método utiliza, como medida de homogeneidade a soma total dos quadrados de uma análise de variância, calculada para cada uma das variáveis. Este método resume-se aos seguintes passos: a) calcula-se as médias das variáveis para cada grupo; b) calculada-se a distância Euclidiana Quadrática; c) somam-se as distâncias para todos os indivíduos; e; d) otimiza-se a variância mínima dentro dos grupos.

A função objetivo que se pretende minimizar é também chamada soma dos quadrados dos erros (SQE). No princípio do processo, cada indivíduo constitui um grupo e tem $SQE = 0$. Logo, o passo seguinte faz agrupar os dois indivíduos que provocam um aumento mínimo no

³⁵ É um ponto único representado pela média de todas as variáveis (MALHOTRA, 2001, p. 530).

³⁶ WARD, J. **Hierarchical grouping to optimize an objective function**, Journal of the American Statistical Association, 1963, 58:236-244.

valor da soma dos quadrados dos erros, passando a existir $n - 1$ grupos; estes $n - 1$ grupos são então reexaminados e transformados em $n - 2$ grupos, mas de tal modo que o aumento na função objetivo seja minimizado; o processo continua de forma sistemática até que todos os indivíduos formem um grupo apenas.

Conforme Reis (1997, p. 322), “este processo tende a formar grupos de tamanho semelhante”, contudo, como se busca grupos o mais homogêneos internamente, é preferível aceitar esta possível – e não factível – desvantagem em relação aos métodos de ligação simples e completa. Então, para ilustrar o método, os valores presentes na Tabela 4 serão novamente utilizados.

Embora, o método de Ward consista em agrupar indivíduos com base na soma dos quadrados entre os dois agrupamentos somados sobre todas as observações, tem-se a matriz distância Euclidiana quadrada para agrupar os indivíduos. Ela é descrita como:

$$D_1 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 931 & 515 & 477 & 1842 \\ \dots & 0 & \mathbf{78} & 454 & 4545 \\ \dots & \dots & 0 & 314 & 3559 \\ \dots & \dots & \dots & 0 & 4161 \\ \dots & \dots & \dots & \dots & 0 \end{bmatrix} \end{matrix}$$

A soma do quadrado dos desvios nada mais é do que a observação pela média das n observações e j variáveis. A função objetivo a minimizar é:

$$SQD = \sum_{j=1}^p \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \quad (1.4)$$

Então, têm-se:

$$SQD_{12} = [(20-12,5)^2 + (5-12,5)^2] + [(18-13,5)^2 + (9-13,5)^2] + [(11-23)^2 + (35-23)^2] + [(10-6,5)^2 + (3-6,5)^2] = 465,50;$$

$$SQD_{13} = [(20-15,5)^2 + (11-15,5)^2] + [(18-14)^2 + (10-14)^2] + [(11-20,5)^2 + (30-20,5)^2] + [(10-8,5)^2 + (7-8,5)^2] = 257,50$$

$$SQD_{14} = [(20-13,5)^2 + (7-13,5)^2] + [(18-10)^2 + (2-10)^2] + [(11-8)^2 + (15-8)^2] + [(10-7)^2 + (4-7)^2] = 288,5;$$

$$SQD_{15} = [(20-34,5)^2 + (49-34,5)^2] + [(18-31,5)^2 + (49-31,5)^2] + [(11-9)^2 + (45-9)^2] + [(10-18)^2 + (26-18)^2] = 2148;$$

$$SQD_{23} = [(5-8)^2 + (11-8)^2] + [(9-9,5)^2 + (10-9,5)^2] + [(35-32,5)^2 + (30-32,5)^2] + [(3-5,5)^2 + (7-5,5)^2] = 39,50;$$

$$SQD_{24} = [(5-6)^2 + (7-6)^2] + [(9-5,5)^2 + (10-5,5)^2] + [(35-25)^2 + (15-25)^2] + [(3-3,5)^2 + (7-3,5)^2] = 247;$$

$$SQD_{25} = [(5-27)^2 + (49-27)^2] + [(9-27)^2 + (45-27)^2] + [(35-21)^2 + (7-21)^2] + [(3-14,5)^2 + (26-14,5)^2] = 2272,50;$$

$$SQD_{34} = [(11-9)^2 + (7-9)^2] + [(10-6)^2 + (2-6)^2] + [(30-22,5)^2 + (15-22,5)^2] + [(7-5,5)^2 + (4-5,5)^2] = 157;$$

$$SQD_{35} = [(11-30)^2 + (49-30)^2] + [(10-27,5)^2 + (45-27,5)^2] + [(30-18,5)^2 + (7-18,5)^2] + [(7-16,5)^2 + (26-16,5)^2] = 1780;$$

$$SQD_{45} = [(7-28)^2 + (49-28)^2] + [(2-23,5)^2 + (45-23,5)^2] + [(15-11)^2 + (7-11)^2] + [(4-15)^2 + (26-15)^2] = 2080,50;$$

PASSO	GRUPOS POSSÍVEIS	SQD	GRUPOS FORMADOS
1°	(1) (2) (3) (4) (5)	0	(1) (2) (3) (4) (5)
2°	(1 2)	465,50	(2 3) (1) (4) (5)
	(1 3)	257,50	
	(1 4)	288,50	
	(1 5)	2148,50	
	(2 3)	39,50	
	(2 4)	247,00	
	(2 5)	2272,50	
	(3 4)	157,00	
	(3 5)	1780,00	
	(4 5)	2080,50	

Quadro 2 – Processo de agrupamento segundo o método de Ward – Passo 1-2.

Observando os valores de SQD no Quadro 2, procura-se pela menor, ou seja, entre os indivíduos 2 e 3. Assim, esses formam um grupo e, os demais, continuam cada um como um grupo. Logo, recalculam-se as SQD para os novos indivíduos:

$$SQD_{23} = [(20-12)^2 + (5-12)^2 + (11-12)^2] + [(18-12,33)^2 + (9-12,33)^2 + (10-12,33)^2] + [(11-25,33)^2 + (35-25,33)^2 + (30-25,33)^2] + [(10-20)^2 + (3-20)^2 + (7-20)^2] = 1041,34;$$

$$SQD_{234} = [(5-7,62)^2 + (11-7,62)^2 + (7-7,62)^2] + [(5-7)^2 + (11-7)^2 + (7-7)^2] + [(5-26,67)^2 + (11-26,67)^2 + (7-26,67)^2] + [(5-4,67)^2 + (11-4,67)^2 + (7-4,67)^2] = 1187,21;$$

$$SQD_{235} = [(5-21,67)^2 + (11-21,67)^2 + (49-21,67)^2] + [(5-21,33)^2 + (11-21,33)^2 + (49-21,33)^2] + [(5-24)^2 + (11-24)^2 + (49-24)^2] + [(5-12)^2 + (11-12)^2 + (49-12)^2] = 5851,68;$$

PASSO	GRUPOS POSSÍVEIS	SQD	GRUPOS FORMADOS
3°	(1 2 3)	1041,34	(1 4) (2 3) (5)
	(1 4)	288,50	
	(1 5)	2148,50	
	(2 3 4)	1187,21	
	(2 3 5)	5851,68	
	(4 5)	2080,50	

Quadro 3 – Processo de agrupamento segundo o método de Ward – Passo 3.

Após 2 e 3 terem formado um grupo e, assim, recalculado os SQD quem minimiza o impacto é entre os indivíduos 1 e 4. Novamente recalculam-se as SQD:

$$SQD_{145} = [(20-25,33)^2 + (7-25,33)^2 + (49-25,33)^2] + [(18-21,67)^2 + (2-21,67)^2 + (45-21,67)^2] + [(11-11)^2 + (15-11)^2 + (7-11)^2] + [(10-40)^2 + (4-40)^2 + (26-40)^2] = 4119,98;$$

$$SQD_{1423} = [(20-10,75)^2 + (7-10,75)^2 + (5-10,75)^2 + (11-10,75)^2] + [(18-9,75)^2 + (2-9,75)^2 + (9-9,75)^2 + (10-9,75)^2] + [(11-22,75)^2 + (15-22,75)^2 + (35-22,75)^2 + (30-22,75)^2] + [(10-6)^2 + (4-6)^2 + (3-6)^2 + (7-6)^2] = 692,22;$$

PASSO	GRUPOS POSSÍVEIS	SQD	GRUPOS FORMADOS
4°	(1 4 5)	4119,98	(1 4 2 3) (5)
	(2 3 5)	5851,68	
	(1 4 2 3)	692,22	

Quadro 4 – Processo de agrupamento segundo o método de Ward – Passo 4.

Quem minimiza a função está ente os grupos já formados (1 4) e (2 3). Unindo-se formam (1 4 2 3). Recalculando analogamente:

$$SQD_{14235} = [(20-18,40)^2 + (7-18,40)^2 + (5-18,40)^2 + (11-18,40)^2 + (49-18,40)^2] + [(18-16,80)^2 + (2-16,80)^2 + (9-16,80)^2 + (10-16,80)^2 + (45-16,80)^2] + [(11-19,60)^2 + (15-19,60)^2 + (35-19,60)^2 + (30-19,60)^2 + (7-19,60)^2] + [(10-6)^2 + (4-10)^2 + (3-10)^2 + (7-10)^2 + (26-10)^2] = 3391,20.$$

PASSO	GRUPOS POSSÍVEIS	SQD	GRUPOS FORMADOS
5°	(1 4 2 3 5)	3391,20	(1 4 2 3 5)

Quadro 5 – Processo de agrupamento segundo o método de Ward – Passo 5.

Tendo sido calculados os valores para a soma do quadrado dos desvios e seguindo os passos supra-expostos têm-se um único grupo final, que reúne na totalidade todos os indivíduos da amostra/exemplo.

Assim, com a aplicação do método em que, no primeiro passo, cada grupo é formado por um único indivíduo; no segundo passo, os indivíduos 2 e 3 formam um grupo por serem aqueles que minimizam o valor da função objetivo. No passo seguinte, o número de grupos reduz-se a três, constituídos por (2 3), (1 4) e (5). No quarto passo, unem-se os dois primeiros grupos formados (1 4 2 3). Por fim, quando a função objetivo atinge o valor de 3391,20, os cinco indivíduos agrupam-se todos, formando apenas um grupo que contém a totalidade dos indivíduos. Fica melhor explicitado com a ajuda do dendrograma da Figura 4.

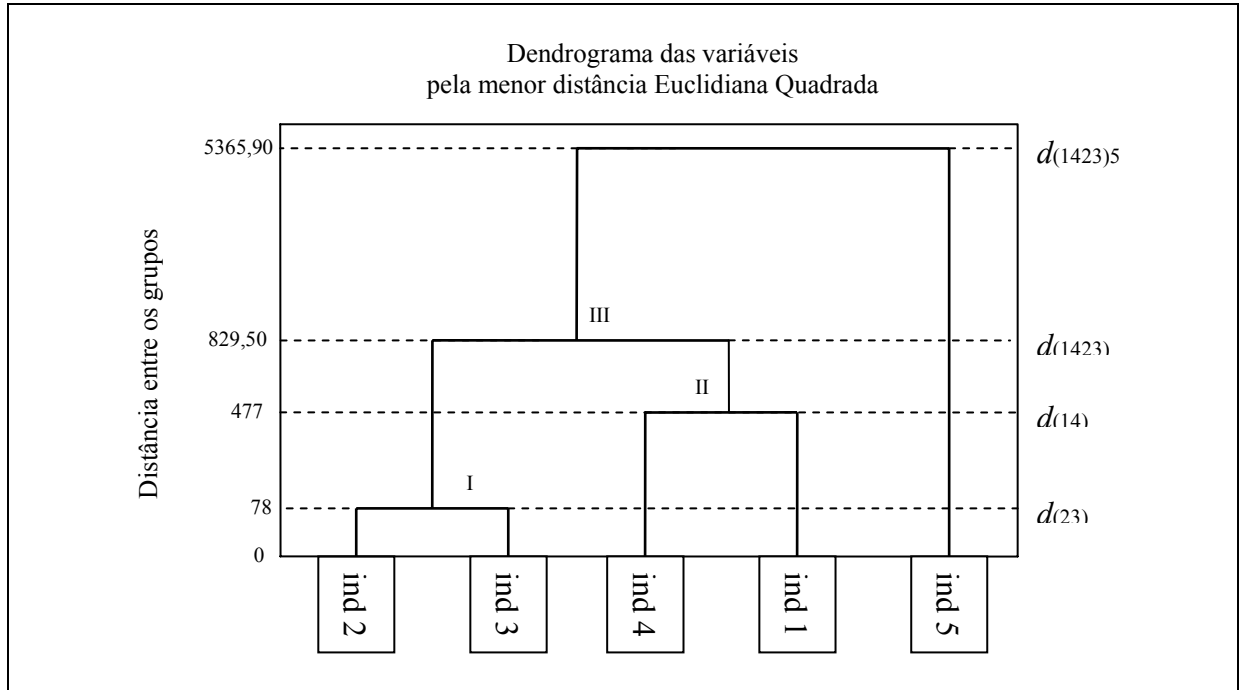


Figura 4 – Dendrograma referente ao agrupamento pelo método de Ward.

4.1.3 Escolha do método e número de grupos

Existem diversas formas para escolher o melhor método de agrupamento. Souza & Vicini (2006, 52-58) apresentam o coeficiente de correlação cofenética. Esse mede a correlação da matriz das distâncias – fenética – com a matriz resultante da simplificação proporcionada pelo método de agrupamento – cofenética. Contudo, com um elevado número de variáveis e indivíduos não existindo um *software* como suplemento para isso a matriz cofenética se torna inviável de construir manualmente, portanto, essa hipótese será descartada neste trabalho.

Mas, estando em acordo com Mingoti (2005) em que, a partição desejada é aquela que produz grupos os mais heterogêneos possíveis e de forma que os indivíduos do interior de cada grupo sejam homogêneos, o método de Ward atendeu tais prerrogativas se comparado com os demais métodos utilizados para a AA.

Quanto ao número de grupos, ainda conforme Mingoti (2005, p. 182), busca-se no gráfico das distâncias de fusão entre os indivíduos o ponto em que ocorre o maior salto. Este ponto indica o número de grupos e a partição ideal ou o momento de parada do algoritmo de agrupamento.

Em outras palavras, à medida que se avança no algoritmo de agrupamento, a similaridade entre os grupos formados que estão sendo combinados nos respectivos, vai decrescendo. E, conseqüentemente, a distância entre eles vai aumentando. Sendo assim, saltos demasiadamente grandes representam uma perda acentuada de similaridade nos grupos formados.

À medida que se passa do estágio k para o estágio $k + 1$ no algoritmo de agrupamento, ou seja, se passa de $n - k$ para $n - k - 1$ a qualidade da partição tende a decrescer. À medida que o nível de fusão aumenta o nível de similaridade decresce. Então, dar “o corte” no dendrograma significa traçar a Linha Fenon³⁷ que delimitará o número de grupos a se formarem. Além de buscar o menor nível de distorção que o processo aglomerativo possa trazer, é preciso ter intuição. Ainda mais, depende não só do maior salto, mas, sobretudo, do conhecimento acumulado pelo pesquisador de seu objeto de pesquisa ao decidir o ponto a ser efetuado “o corte”.

4.2 Análise fatorial de componentes principais

Efetuada os agrupamentos dos municípios, resta agora explicar a relação desses com as 17 variáveis escolhidas para medir o nível de modernização da agricultura nos COREDEs essencialmente agrícolas do Estado do Rio Grande do Sul (RS). De pronto, parece muito fácil explicar a relação entre os grupos mais homogêneos com as variáveis, mas isto não é tão simples visto o elevado número de variáveis. Sendo assim, esta seção se dedica a reduzir o nível de variáveis perdendo o mínimo possível de informação da matriz original de dados.

Então, ainda dentro da análise multivariada, existe mais um instrumento que vem a facilitar a vida do pesquisador, a Análise Fatorial (AF) ou *Factor Analysis*. Assim como no processo de agrupamento, a AF exige a escolha de um algoritmo para extração dos fatores. No presente trabalho fez-se a opção pelo método da Análise de Componentes Principais (ACP).

³⁷ Linha Fenon ou linha de corte que marca a distância onde ocorre o maior salto no gráfico das distâncias de fusão.

4.2.1 Análise de Componentes Principais (ACP)

A Análise das Componentes Principais objetiva descrever os dados contidos num quadro indivíduos-variáveis. Por isso, é considerado um método fatorial, pois a redução do número de variáveis não se faz por simples seleção de algumas variáveis, mas pela construção de novas variáveis, os fatores, obtidos através da combinação linear das variáveis iniciais (BOUROCHE & SAPORTA, 1982).

A ACP fornece medidas responsáveis pelas maiores variações entre os resultados sem que se percam muitas informações ao transformar o conjunto original de variáveis em outro conjunto: os Componentes Principais (CP) de dimensões equivalentes. Tal transformação ocorre com a menor perda de informação possível, sendo que essa também pode auxiliar a eliminar algumas variáveis originais, no caso de possuírem pouca informação.

A Análise das Componentes Principais aborda aspectos como geração, seleção e interpretação das componentes investigadas e, ainda, determina as variáveis de maior influência na formação de cada componente. Conforme Regazzi (2001, p. 1), “procura-se redistribuir a variação nas variáveis (eixos originais) de forma a obter o conjunto ortogonal de eixos não correlacionados”. Então, um grande número de informações sobre indivíduos, como é o caso do presente estudo, é transformado em um novo conjunto de variáveis que mantém, ao máximo, a variabilidade original do conjunto.

A ACP permite agrupar indivíduos similares mediante inspeções visuais em dispersões gráficas no espaço bi ou tridimensional, o que facilita a interpretação geométrica. Esta redução da dimensionalidade é chamada transformação de *Karhunen-Loève*, ou, Análise de Componentes Principais, no qual os autovalores encontrados são denominados de componentes principais.

O algoritmo se baseia na matriz de variância-covariância ou na matriz de correlação, de onde são extraídos os autovalores e os autovetores. Em suma, o que o método traduz, é a criação de novas variáveis não correlacionadas, obtidas de combinações lineares das variáveis iniciais que são apresentadas em ordem decrescente relativas ao seu poder de explicação – as componentes principais.

O processo tem cinco passos: a) coleta de dados para a matriz (X_n); b) determinação da matriz de correlação (R), ou a matriz de variância-covariância amostral (S) dos dados

originais (X_n); c) estimar os autovalores ($\hat{\Lambda}$) – raízes características; e, por fim, d) estimar os autovetores (\vec{x}) que possibilitam encontrar as novas variáveis (Y_n).

De pronto, calcula-se a matriz S, ou matriz R, e verifica-se se as variáveis estão correlacionadas uma com as outras. Recomenda-se aplicar o teste *Kaiser-Meyer-Olkin Measure of Adequacy* (KMO) que compara a magnitude dos coeficientes de correlação parcial e varia entre 0 e 1, sendo assim calculado:

$$KMO = \frac{\sum_{i \neq j} \sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} \sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} \sum_{i \neq j} a_{ij}^2} \quad (1.5)$$

em que r_{ij} é o coeficiente de correlação simples entre as variáveis x_i e x_j e a_{ij} é o coeficiente de correlação parcial entre as mesmas. Pequenos valores para o KMO indicam a não adequabilidade dos dados para o processo. Assim, a correlação entre os pares de variáveis não pode ser explicada por outras variáveis. O Tabela 6 sintetiza a interpretação do teste KMO:

Tabela 6 – Resultado para o teste KMO e a adequabilidade dos dados.

KMO	Análise das componentes principais
1 – 0,90	Muito boa
0,80 – 0,90	Boa
0,70 – 0,80	Média
0,60 – 0,70	Razoável
0,50 – 0,60	Má
KMO < 0,50	Inaceitável

Fonte: Hair Jr. (2005).

Juntamente com KMO, Reis (1997, p. 278) aponta outro teste que coaduna para checar a validade da aplicação deste tipo de análise. O Teste de esfericidade de Bartlett testa a hipótese de a matriz de correlações ser uma matriz identidade³⁸ e seu determinante ser igual a 1, logo, as variáveis não estariam correlacionadas. Então, para testar o grau de adequação dos dados ao método, deve-se verificar se a matriz R é ou não uma matriz identidade. Então, testa-

³⁸ Uma matriz identidade é toda matriz quadrada com diagonal principal 1 e os demais elementos são nulos. Na ocorrência de se aceitar a hipótese nula, a existência de uma matriz identidade, ou seja, se reconhece a falta de correlação entre as variáveis, o que invalidaria o prosseguimento do método por falta de adequação dos dados aos pressupostos do método.

se a hipótese nula contra a alternativa, ou seja, a hipótese nula³⁹ ($H_0 : P = 1$) ou hipótese alternativa ($H_1 : P \neq 1$). Dessa forma, para aplicação do método, espera-se que H_0 seja rejeitada.

A matriz variância-covariância é expressa pelas ligações realizadas entre as p variáveis, tomadas duas a duas resumidas por suas covariâncias S_{ij} . Conforme Johnson & Wichern (1992), a matriz quadrada, simétrica, denota-se por:

$$S = \begin{bmatrix} \hat{V}ar(X_1) & \hat{C}ov(X_1, X_2) & \dots & \hat{C}ov(X_1, X_p) \\ \hat{C}ov(X_1, X_2) & \hat{V}ar(X_2) & \dots & \hat{C}ov(X_2, X_p) \\ \dots & \dots & \ddots & \dots \\ \hat{C}ov(X_1, X_p) & \hat{C}ov(X_2, X_p) & \dots & \hat{V}ar(X_p) \end{bmatrix} \quad (1.6)$$

Sendo S , o conjunto de variância-covariância representado pela matriz S , denominada matriz de variância-covariância amostral das p variáveis. Assim, o termo situado na intersecção da i -ésima linha da j -ésima coluna é a covariância (S_{ij}) e os termos da diagonal principal são as variâncias (S_i^2); sendo definidos, respectivamente, por:

$$\hat{C}ov(X_j, X_j) = \frac{1}{n-1} \left[\sum_{i=1}^n X_{ij} X_{ij} - \frac{\left(\sum_{i=1}^n X_{ij} \right) \left(\sum_{i=1}^n X_{ij} \right)}{n} \right] \quad (1.7)$$

$$\hat{V}ar(X_j) = \frac{1}{n-1} \left[\sum_{i=1}^n X_{ij}^2 - \frac{\left(\sum_{i=1}^n X_{ij} \right)^2}{n} \right] \quad (1.8)$$

Considere os dados contidos na Tabela 7, referentes a cinco indivíduos e com duas variáveis referentes aos mesmos. As componentes principais serão calculados a partir da matriz de variância-covariância descrita anteriormente.

³⁹ A notação “P=1” é meramente um indicativo da presença/ocorrência de matriz identidade da matriz R .

Tabela 7 – Observações relativas às duas variáveis, levantadas em cinco indivíduos.

Observações	Variável X_1	Variável X_2
1	62	61
2	71	58
3	69	94
4	55	35
5	50	18

Fonte: Elaboração própria.

Anterior as matrizes de correlação $R_{2 \times 2}$ e matriz de variância-covariância $S_{2 \times 2}$ faz-se necessário uma análise descritiva das duas variáveis observadas:

Tabela 8 – Estatística descritiva relativa às duas variáveis, levantadas em cinco indivíduos.

	Variável X_1	Variável X_2
Média aritmética das variáveis	61,40	53,20
Somatório das variáveis	307	266
Somatório do quadrado de X_i	19171	17470
Somatório do produto das variáveis		17211
Variância amostral das variáveis	80,30	829,70
Desvio padrão amostral das variáveis	8,96	28,80

Fonte: Elaboração própria.

A matriz variância⁴⁰ e covariância S são geradas pelas expressões 1.8 e 1.7 respectivamente. Para $p=2$ variáveis e $n=5$ indivíduos, têm-se:

$$\hat{V}ar(X_1) = \frac{1}{5-1} \left[19171 - \frac{307^2}{5} \right] = 80,30$$

$$\hat{V}ar(X_2) = \frac{1}{5-1} \left[17470 - \frac{266^2}{5} \right] = 829,70$$

$$\hat{C}ov(X_1, X_2) = \frac{1}{5-1} \left[17211 - \frac{307 \times 266}{5} \right] = 219,65$$

Logo, chega-se a matriz $S_{2 \times 2}$:

⁴⁰ A variância já está presente na análise descritiva e na Tabela 8, contudo, cabe demonstrar seu cálculo.

$$S = \begin{bmatrix} 80,30 & 219,65 \\ 219,65 & 829,70 \end{bmatrix}.$$

Quanto a matriz de correlação, Rummel (1967, p. 460) salienta que “é com maior frequência empregada em técnicas de análise fatorial”. A matriz de correlação é utilizada no caso de os dados necessitarem de uma padronização (ZAMBRANO & LIMA, 2004, p. 557). É válido, mais uma vez, lembrar que os valores originais sejam padronizados/normalizados para se obter a matriz R . Contudo, o *software* *Statística 6.0* executa esta tarefa. A normalização consiste em expressar, em desvios-padrão, os desvios das observações originais em relação à sua média aritmética. Assim, sejam n observações e X_{ij} a i -ésima observação da j -ésima variável, a correspondente variável Z_{ij} , normalizada, é obtida por:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{\sqrt{\text{Var}X_j}} \quad (i=1, 2, \dots, n) \quad (1.9)$$

A matriz R é denotada por:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \dots & \dots & \ddots & \dots \\ r_{p1} & r_{p2} & \dots & r_{pp} \end{bmatrix} \quad (1.10)$$

onde:

$$r_{pp} = r(X_j, X_j) = \frac{\text{Cov}(X_j, X_j)}{\sqrt{\text{Var}(X_j) \cdot \text{Var}(X_j)}} \quad (1.11)$$

A matriz R , que é extraída das variáveis originais – se extraída das normalizadas o valor será o mesmo –, será calculada por:

$$r_{x_1x_2} = \frac{219,65}{\sqrt{80,30 \cdot 829,70}} = 0,85.$$

A correlação entre a variável, em relação a ela mesma, é fornecida por:

$$r_{11} = \frac{\text{Cov}(X_1, X_1)}{S_{x_1} \cdot S_{x_1}} = \frac{S_{x_1}^2}{S_{x_1}^2};$$

$$r_{11} = \frac{80,30^2}{80,30^2} = 1;$$

logo, a matriz R será:

$$R = \begin{bmatrix} 1 & 0,85 \\ 0,85 & 1 \end{bmatrix}.$$

Padronizando-se as variáveis:

$$Z_{11} = \frac{62 - 61,40}{\sqrt{80,30}} = 0,07$$

$$Z_{21} = \frac{61 - 53,20}{\sqrt{829,70}} = 0,27$$

$$Z_{12} = \frac{71 - 61,40}{\sqrt{80,30}} = 1,07$$

$$Z_{22} = \frac{58 - 53,20}{\sqrt{829,70}} = 0,17$$

$$Z_{13} = \frac{69 - 61,40}{\sqrt{80,30}} = 0,85$$

$$Z_{23} = \frac{94 - 53,20}{\sqrt{829,70}} = 1,42$$

$$Z_{14} = \frac{55 - 61,40}{\sqrt{80,30}} = -0,71$$

$$Z_{24} = \frac{35 - 53,20}{\sqrt{829,70}} = -0,63$$

$$Z_{15} = \frac{50 - 61,40}{\sqrt{83,20}} = -1,25$$

$$Z_{25} = \frac{18 - 53,20}{\sqrt{829,70}} = -1,22$$

A solução pela matriz de correlação é recomendada quando as variáveis são medidas em escalas diferentes entre si, pois essa matriz é equivalente a matriz das variáveis padronizadas (JOHNSON & WICHERN, 1992).

Seguindo o raciocínio, restam achar os autovalores e autovetores. O escalar $\hat{\Lambda}$ será chamado de autovalor – *eigenvalues* – e o vetor \vec{x} um autovetor. Em um primeiro momento, parte-se de S , e I (matriz identidade $p_x p$), então os escalares $\hat{\Lambda}_1, \hat{\Lambda}_2, \dots, \hat{\Lambda}_p$ satisfazem a equação polinomial que segue:

$$|S - \hat{\Lambda} I| = 0 \tag{1.12}$$

Seja S a matriz de variância-covariância de dimensão $p_x p$, e seja $\hat{\Lambda}$ um autovalor de S . Logo \vec{x} é um vetor não nulo, tal que:

$$S\vec{X} = \hat{\Lambda}\vec{X}, \quad (1.13)$$

no qual, \vec{X} é uma matriz $p \times p$ de todos autovetores e $\hat{\Lambda}$ é um escalar. Assim, \vec{x} é tido como um autovetor ou vetor característico da matriz S , associada com o valor de $\hat{\Lambda}$.

Para determinar as componentes principais, partindo da matriz S , resolve-se a equação 1.12. Substituindo-se a equação pelas matrizes S e I , obtém-se:

$$\left[\begin{array}{cc} 80,30 & 219,65 \\ 219,65 & 829,70 \end{array} \right] - \hat{\Lambda} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

Multiplicando o autovalor $\hat{\Lambda}$ pela matriz I , têm-se as seguintes matrizes:

$$\left[\begin{array}{cc} 80,30 & 219,65 \\ 219,65 & 829,70 \end{array} \right] - \begin{bmatrix} \hat{\Lambda} & 0 \\ 0 & \hat{\Lambda} \end{bmatrix} = 0$$

Subtraindo uma matriz pela outra, obtém-se:

$$\left[\begin{array}{cc} 80,30 - \hat{\Lambda} & 219,65 \\ 219,65 & 829,70 - \hat{\Lambda} \end{array} \right] = 0$$

O próximo passo é calcular o determinante da matriz. Chega-se ao seguinte resultado que corresponde aos autovalores da matriz S :

$$\begin{aligned} (80,30 - \hat{\Lambda})(829,70 - \hat{\Lambda}) - (219,65)^2 &= 66.624,91 - 80,30\hat{\Lambda} - 829,70\hat{\Lambda} + \hat{\Lambda}^2 - 48.246,12 \\ &= \hat{\Lambda}^2 - 910\hat{\Lambda} + 18.378,79 = 0 \end{aligned}$$

Os autovalores – raízes características – são obtidos da seguinte equação:

$$\hat{\Lambda} = \frac{-b \pm \sqrt{(b)^2 - 4(a)(c)}}{2.(a)} \quad (1.14)$$

então:

$\hat{\Lambda} = \frac{910 \pm \sqrt{(-910)^2 - 4(1)(18.378,79)}}{2 \cdot (1)}$, logo, os dois autovalores resultantes da equação são:

$$\hat{\Lambda}_1 = 889,33 \text{ e } \hat{\Lambda}_2 = 20,67.$$

Através dos autovalores obtidos deve-se chegar ao determinante da matriz S e ao traço da matriz:

$$\sum_{j=1}^k \hat{\Lambda}_j = \hat{\Lambda}_1 + \hat{\Lambda}_2 + \dots + \hat{\Lambda}_p \quad (1.15)$$

Ou seja:

$$889,33 + 20,67 = 910 \Rightarrow \text{traço da matriz } S.$$

O determinante é dado por:

$$(\hat{\Lambda}_1) \cdot (\hat{\Lambda}_2) \cdot \dots \cdot (\hat{\Lambda}_p) \quad (1.16)$$

Então:

$$(889,33)(20,67) = 18382,45 \Rightarrow \text{determinante da matriz } S.$$

Lembrando que, a contribuição de cada componente principal Y_i é medida em termos de variância. Logo, tem-se que o quociente é expresso em percentagem:

$$\frac{V\hat{a}r(Y_p)}{\sum_{i=1}^p V\hat{a}r(Y_p)} \cdot 100 = \frac{\hat{\Lambda}_p}{\sum_{j=1}^p \hat{\Lambda}_j} \cdot 100 = \frac{\hat{\Lambda}_p}{\text{traço}S} \cdot 100, \quad (1.17)$$

sendo que, esta expressão representa a proporção da variância total explicada pelas n componentes encontradas. Surge uma questão inerente de quais componentes utilizar no estudo. Assim, selecionam-se as componentes que acumulam uma percentagem de variância explicada, igual ou superior a 70%. Ou seja, fica-se com Y_1, \dots, Y_k tal que:

$$\frac{V\hat{a}r(Y_1) + \dots + V\hat{a}r(Y_k)}{\sum_{i=1}^p V\hat{a}r(Y_i)} \cdot 100 \geq 70\% \text{ no qual, } k < p. \quad (1.18)$$

O sucesso da metodologia será medido pelo valor de k . Se $k=1$, significa que o método está reduzido ao máximo à dimensão original. Assim, podem-se comparar os indivíduos em uma escala linear. Caso $k=2$ é possível localizar cada indivíduo em um plano cartesiano, sendo que dois eixos representam as duas componentes encontradas, e suas coordenadas formam uma elipse. Se k for maior que dois, complica-se a comparação (REGAZZI, 2001).

No exemplo até aqui desenvolvido, têm-se dois eixos perpendiculares, cujas coordenadas estão representadas pelos autovetores I e II da matriz $S(\bar{x}_1, \bar{x}_2)$. Os elementos desses vetores definem sua posição, isto é, o ângulo que eles formam com os eixos originais de Y_1 e Y_2 . O comprimento dos vetores corresponde aos autovalores ($\hat{\Lambda}_1 = 889,33$, $\hat{\Lambda}_2 = 20,67$) que representam a variância dos novos eixos. A Figura 5 representa os autovalores e autovetores.

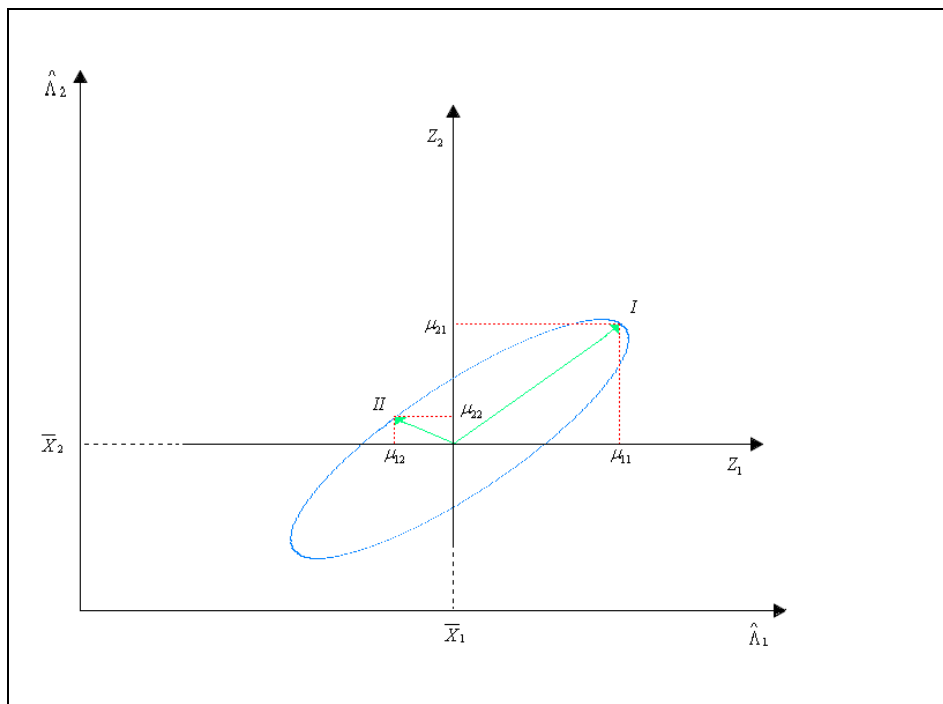


Figura 5 – Representação gráfica dos autovalores e autovetores.

A proporção da variância total explicada por cada componente principal é dada por:

$$\frac{\hat{\Lambda}_p}{\text{traço}S} \cdot 100 \quad (1.19)$$

Como sabido, os autovalores estão em ordem decrescente segundo seus valores. Logo, a explicação de cada componente principal será, respectivamente:

$$\frac{889,33}{910} \cdot 100 = 97,73\% \Rightarrow \hat{\Lambda}_1;$$

$$\frac{20,67}{910} \cdot 100 = 2,27\% \Rightarrow \hat{\Lambda}_2.$$

Então, tem-se que, a variância distribuída entre $\hat{\Lambda}_1 = 889,33$ e $\hat{\Lambda}_2 = 20,67$ está explicada 97,73% pelo primeiro eixo fatorial e 2,27% pelo segundo. Sendo assim, cada componente principal sintetiza a máxima proporção de variância contida nos dados da matriz de dados. Aqui, no exemplo, percebe-se que apenas uma componente principal satisfaz as exigências de 1.18 – $\hat{\Lambda}_1 = 889,33$ – contudo, relaxa-se a hipótese e prossegue-se com o exemplo/raciocínio até aqui desenvolvido.

Sejam $\hat{\Lambda}_1, \hat{\Lambda}_2, \dots, \hat{\Lambda}_p$ as p soluções, tem-se que cada autovalor $\hat{\Lambda}_j$ corresponde a um autovetor característico.

Sendo assim, as coordenadas de x_{11} e x_{12} do autovetor \overline{X}_1 são calculadas pela equação matricial:

$$|S - \hat{\Lambda}_j I| \overline{X}_j = 0 \quad (1.20)$$

Assim, para o primeiro autovalor:

$$|S - \hat{\Lambda}_1 I| \overline{X}_1 = 0$$

O cálculo da primeira componente principal referente a $\hat{\Lambda}_1 = 889,33$ será dado pelo autovetor associado a $\hat{\Lambda}_1$, sendo que a equação característica dos autovetores é

$|S - \hat{\Lambda}_1 I| \overline{X}_1 = 0$. Existe um vetor \vec{x} para cada valor de $\hat{\Lambda}$. Pois, $\vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix}$ é um vetor não

normalizado. Em que \vec{o} é um vetor nulo, de dimensão $p_x 1$.

Substituindo-se pelas matrizes S, I , pelo primeiro autovetor e pela matriz de incógnitas \vec{x}_i , tem-se:

$$\begin{bmatrix} 80,30 & 219,65 \\ 219,65 & 829,70 \end{bmatrix} - 889,33 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Multiplicando-se o autovalor $\hat{\Lambda}_1$ pela matriz I e subtraindo da matriz S , chega-se as seguintes matrizes:

$$\begin{bmatrix} 80,30 - 889,33 & 219,65 \\ 219,65 & 829,70 - 889,33 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Tais matrizes dão origem ao seguinte sistema:

$$\begin{cases} -809,03x_{11} + 219,65x_{12} = 0 \\ 219,65x_{11} - 59,63x_{12} = 0 \end{cases}.$$

Tal sistema de equações é indeterminado, em virtude de $|S - \hat{\Lambda} I| = 0$. Sendo assim, parte-se para:

$$\begin{bmatrix} -809,03 & 219,65 \\ 219,65 & -59,63 \end{bmatrix} = 0;$$

ou, equivalendo dizer que quando o vetor passa pela origem, fica $x_{11} = x_{12} = 0$.

Devido a isto, pode-se deixar uma das equações – a segunda –, e atribuir um valor arbitrário, não nulo, a uma das incógnitas – $x_{12} = 1$. Dessa forma, tem-se que:

$$-809,03x_{11} + 219,65(1) = 0$$

$$-809,03x_{11} = -219,65, \text{ logo o valor da incógnita } x_{11} \text{ será } 0,27.$$

O autovetor associado ao primeiro $\hat{\Lambda}_1 = 889,33$ será:

$$\vec{x}_1 = \begin{bmatrix} 0,27 \\ 1 \end{bmatrix} \text{ e sua norma será dada por:}$$

$$\|\vec{x}_1\| = \sqrt{(0,27)^2 + (1)^2} = 1,04.$$

Hoffmann (1999, p. 5) lembra de duas restrições quanto à normalidade e ortogonalidade:

$$\text{a) } \vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix} \text{ com } \sum_{j=1}^p x_{ij}^2 = 1 \text{ sob a condição de normalidade } (\vec{x}_i \cdot \vec{x}_i = 1) \text{ e, também,}$$

b) $\sum_{j=1}^p x_{ij} x_{kj} = 0$ para $i \neq k$ ($\vec{x}_i \cdot \vec{x}_k = 0$ para $i \neq k$), sendo esta a condição de ortogonalidade dos vetores.

A normalidade é a primeira restrição para que o sistema tenha solução única, e a ortogonalidade garante que as componentes principais sejam independentes. Melhor dizendo, significa que cada vetor é normalizado, ou seja, a soma dos quadrados dos coeficientes é igual a 1, sendo, ainda, ortogonais entre si.

O autovetor normalizado é dado por:

$$\vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix} = \frac{1}{\sqrt{x_{i1}^2 + x_{i2}^2 + \dots + x_{ip}^2}} \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix} = \frac{\vec{x}_i}{\|\vec{x}_i\|} \cdot \vec{x}_i = 1 \quad (1.21)$$

Para que esse vetor seja unitário, é necessário normalizar o autovetor a 1, da seguinte maneira:

$$x_i = \frac{1}{\|\vec{x}_i\|} \vec{x}_i \quad (1.22)$$

No exemplo, substituindo em 1.22 pelos seus respectivos valores tem-se:

$$x_1 = \frac{1}{\|\vec{x}_1\|} \cdot \vec{x}_1 = x_1 = \frac{1}{\|1,04\|} \cdot \begin{bmatrix} 0,27 \\ 1 \end{bmatrix},$$

logo, o primeiro autovetor normalizado será:

$$x_1 = \begin{bmatrix} 0,26 \\ 0,96 \end{bmatrix},$$

e a sua norma será:

$$\|x_1\| = \sqrt{(0,26)^2 + (0,96)^2} = 1.$$

Conforme Regazzi (2001), tomando os elementos relativos ao vetor \vec{x}_i , assim determinam-se os coeficientes de Y_i , tendo que o i -ésimo componente principal é dado por:

$$Y_i = x_{i1}X_1 + \dots + x_{i2}X_2 + \dots + x_{ip}X_p \quad (1.23)$$

Têm-se ainda:

$$a) \hat{V}ar(Y_1) = \hat{\Lambda}_1 \text{ logo } \hat{V}ar(Y_1) > \hat{V}ar(Y_2) > \dots > \hat{V}ar(Y_p);$$

$$b) \sum \hat{V}ar(X_1) = \sum \hat{\Lambda}_1 = \sum \hat{V}ar(Y_1); \text{ e, } C\hat{o}v(Y_i, Y_j) = 0, \text{ desde que } \sum_{j=1}^p x_{ij}x_{kj} = 0.$$

Assim sendo, a primeira componente principal equivale a: $Y_1 = 0,26X_1 + 0,96X_2$.

Resta encontrar a segunda componente principal, conforme a expressão 1.20, que é oriunda da raiz $\hat{\Lambda}_2 = 20,67$, dado que:

$$|S - \hat{\Lambda}_2 I| \vec{X}_2 = 0.$$

Novamente, devem-se substituir na equação as matrizes S, I , pelo segundo autovetor $\hat{\Lambda}_2 = 20,67$, seguido da matriz de incógnitas (\vec{x}), obtendo-se:

$$\begin{bmatrix} 80,30 & 219,65 \\ 219,65 & 829,70 \end{bmatrix} - 20,67 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Multiplicando-se o autovalor $\hat{\Lambda}_2$ pela matriz I e subtraindo da matriz S , chega-se a seguinte expressão:

$$\begin{bmatrix} 80,30 - 20,67 & 219,65 \\ 219,65 & 829,70 - 20,67 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Tais matrizes dão origem ao seguinte sistema:

$$\begin{cases} 59,63x_{21} + 219,65x_{22} = 0 \\ 219,65x_{21} + 809,03x_{22} = 0 \end{cases}.$$

Tal sistema de equações é indeterminado, em virtude de $|S - \hat{\Lambda}I| = 0$, sendo assim, parte-se para:

$$\begin{bmatrix} 59,63 & 219,65 \\ 219,65 & 809,03 \end{bmatrix} = 0;$$

ou, equivalendo dizer que quando o vetor passa pela origem, equivale a $x_{21} = x_{22} = 0$.

Novamente, devido a isto, deixa-se uma das equações – a segunda –, e atribui-se um valor arbitrário, não nulo, a uma das incógnitas – $x_{22} = 1$. Dessa forma, tem-se que:

$$59,63x_{21} + 219,65(1) = 0$$

$$59,63x_{21} = -219,65, \text{ logo o valor da incógnita } x_{21} \text{ será } -3,68.$$

O autovetor associado ao segundo $\hat{\Lambda}_2 = 20,67$ será:

$$\vec{x}_2 = \begin{bmatrix} -3,68 \\ 1 \end{bmatrix} \text{ e sua norma será dada por:}$$

$$\|\vec{x}_2\| = \sqrt{(-3,68)^2 + (1)^2} = 3,82.$$

Para que esse vetor seja unitário, é necessário normalizar o autovetor a 1 – seguindo a expressão 1.22 –, da seguinte maneira:

$$x_2 = \frac{1}{\|\vec{x}_2\|} \cdot \vec{x}_2.$$

Substituindo-se essa expressão pelos seus respectivos valores tem-se:

$$x_2 = \frac{1}{\|3,82\|} \cdot \begin{bmatrix} -3,68 \\ 1 \end{bmatrix},$$

logo, o autovetor normalizado será:

$$x_2 = \begin{bmatrix} -0,96 \\ 0,26 \end{bmatrix},$$

e a sua norma será:

$$\|x_2\| = \sqrt{(-0,96)^2 + (0,26)^2} = 1$$

Assim sendo, a segunda componente principal equivale a: $Y_2 = -0,96X_1 + 0,26X_2$.

Sendo que os dois vetores são ortogonais, pois $\vec{x}_1 \cdot \vec{x}_2 = 0$ – o produto escalar é igual a zero –, equivalendo à segunda restrição feita por Hoffmann (1999, p. 5).

Para que seja satisfeita, deve-se multiplicar o primeiro autovetor normalizado transposto pelo segundo autovetor normalizado, ou seja:

$$x_1' x_2 = [0,26 \quad 0,96] \begin{bmatrix} -0,96 \\ 0,26 \end{bmatrix}.$$

Multiplicando-se os autovalores normalizados, tem-se a seguinte expressão:

$$x_1' x_2 = (0,26)(-0,96) + (0,96)(0,26) = 0$$

Satisfeitas as condições, o conjunto de n observações e de p -variáveis, é possível encontrar as novas variáveis como o mencionado, denominadas de $Y_k, k = 1, \dots, p$, que são as

combinações lineares das variáveis originais, e apresentam um grau de variabilidade diferente umas das outras. Sendo importante ressaltar, em ACP, a unidade de medida se baseia em combinações lineares não correlacionadas, por isso são de difícil interpretação, e também é por esse motivo que as variáveis originais devem estar na mesma unidade de medida, ou, padronizadas. Então:

$$Y_1 = 0,26X_1 + 0,96X_2$$

$$Y_2 = -0,96X_1 + 0,26X_2$$

$$Y_{11} = 0,26(62) + 0,96(61) = 65,32$$

$$Y_{21} = -0,96(62) + 0,41(61) = -34,51$$

$$Y_{12} = 0,26(71) + 0,96(58) = 74,14$$

$$Y_{22} = -0,96(71) + 0,26(58) = -53,08$$

$$Y_{13} = 0,26(69) + 0,96(94) = 108,18$$

$$Y_{23} = -0,96(69) + 0,26(94) = -41,80$$

$$Y_{14} = 0,26(55) + 0,96(35) = 47,90$$

$$Y_{24} = -0,96(55) + 0,26(35) = -43,70$$

$$Y_{15} = 0,26(50) + 0,96(18) = 30,28$$

$$Y_{25} = -0,96(50) + 0,26(18) = -43,32$$

Na Tabela 9 têm-se as observações originais e as novas componentes geradas a partir de combinações lineares:

Tabela 9 – Matriz de dados originais e a nova matriz gerada – MS.

Observações	Variáveis originais		Novas variáveis	
	X ₁	X ₂	Y ₁	Y ₂
1	62	61	65,32	-34,51
2	71	58	74,14	-53,08
3	69	94	108,18	-41,80
4	55	35	47,90	-43,70
5	50	18	30,28	-43,32

Fonte: Elaboração própria.

Efetuada a análise de componentes principais, deve-se identificar, ou mesmo, entender o significado de cada componente no estudo que está sendo realizado. A interpretação de cada componente principal é feita mediante seu grau de contribuição, ou seja, o grau de influência que cada variável tem sobre cada componente, sendo que esta importância é dada pela correlação entre cada variável X_j e a componente Y_i .

Então, para finalizar a análise de componentes principais, é necessário fazer a correlação entre as variáveis X_j e Y_i , como se pode verificar a seguir:

$$r_{x_1 y_1} = \sqrt{\hat{\Lambda}_1} \cdot \frac{x_{11}}{\sqrt{\hat{V}ar(X_1)}}$$

$$r_{x_1 y_1} = \sqrt{889,33} \cdot \frac{0,26}{\sqrt{80,30}} = 0,87;$$

$$r_{x_2 y_1} = \sqrt{\hat{\Lambda}_1} \cdot \frac{x_{12}}{\sqrt{\hat{V}ar(X_2)}}$$

$$r_{x_2 y_1} = \sqrt{889,33} \cdot \frac{0,96}{\sqrt{829,70}} = 0,99;$$

$$r_{x_1 y_2} = \sqrt{\hat{\Lambda}_2} \cdot \frac{x_{21}}{\sqrt{\hat{V}ar(X_1)}}$$

$$r_{x_1 y_2} = \sqrt{20,67} \cdot \frac{-0,96}{\sqrt{80,30}} = -0,49;$$

$$r_{x_2 y_2} = \sqrt{\hat{\Lambda}_2} \cdot \frac{x_{22}}{\sqrt{\hat{V}ar(X_2)}}$$

$$r_{x_2 y_2} = \sqrt{20,67} \cdot \frac{0,26}{\sqrt{829,70}} = 0,04.$$

Observa-se que a correlação da componente Y_1 é a mais significativa, corroborando com as exigências de 1.18.

A Tabela 10 mostra as componentes principais encontradas no exemplo exploratório, os autovalores, os autovetores, a correlação entre as variáveis, a percentagem de explicação de cada componente e a percentagem total da variância acumulada pelas componentes principais calculadas através da matriz de variância-covariância.

Tabela 10 – Resumo da análise de componentes principais pela matriz variância-covariância.

Componentes principais	Autovalores	Autovetores		Correlação entre X_j e Y_i		Percentagem da variância de Y_i	Percentagem acumulada a variância dos Y_i
		X_1	X_2	X_1	X_2		
Y_1	889,33	0,26	0,96	0,87	0,99	97,73	97,73%
Y_2	20,67	-0,96	0,26	-0,49	0,04	2,27	100%

Fonte: Elaboração própria.

Da mesma forma que o exemplo foi desenvolvido até aqui utilizando a matriz S , pode-se efetuar-lo também com a matriz R . Então, para encontrar os autovalores, agora, a partir da matriz R , deve-se partir da equação característica:

$$|R - \hat{\Lambda} I| = 0 \quad (1.24)$$

Substituindo-se a equação pelas matrizes R e I , obtém-se:

$$\begin{vmatrix} 1 & 0,85 \\ 0,85 & 1 \end{vmatrix} - \hat{\Lambda} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 0$$

Multiplicando o autovalor $\hat{\Lambda}$ pela matriz I , têm-se as seguintes matrizes:

$$\begin{vmatrix} 1 & 0,85 \\ 0,85 & 1 \end{vmatrix} - \begin{vmatrix} \hat{\Lambda} & 0 \\ 0 & \hat{\Lambda} \end{vmatrix} = 0$$

Subtraindo uma matriz pela outra, obtém-se:

$$\begin{vmatrix} 1 - \hat{\Lambda} & 0,85 \\ 0,85 & 1 - \hat{\Lambda} \end{vmatrix} = 0$$

Analogamente, calcula-se o determinante desta matriz que originará aos autovalores da matriz R :

$$\begin{aligned} (1 - \hat{\Lambda})(1 - \hat{\Lambda}) - (0,85)^2 &= 1 - \hat{\Lambda} - \hat{\Lambda} + \hat{\Lambda}^2 - 0,72 \\ &= \hat{\Lambda}^2 - 2\hat{\Lambda} + 0,28 = 0 \end{aligned}$$

Os autovalores – raízes características – são obtidos da seguinte equação (1.14). Então:

$$\hat{\Lambda} = \frac{2 \pm \sqrt{(2)^2 - 4(1)(0,28)}}{2 \cdot (1)}, \text{ logo, os dois autovalores resultantes da equação são:}$$

$$\hat{\Lambda}_1 = 1,85 \text{ e } \hat{\Lambda}_2 = 0,15.$$

Através dos autovalores obtidos é possível se chegar ao determinante da matriz R e ao traço da matriz:

$$\sum_{j=1}^p \hat{\Lambda}_j = \hat{\Lambda}_1 + \hat{\Lambda}_2 + \dots + \hat{\Lambda}_p = \text{traço da matriz } R. \text{ Ou seja:}$$

$$1,85 + 0,15 = 2 \Rightarrow \text{traço da matriz } R.$$

O determinante é dado por (1.16):

$$(\hat{\Lambda}_1) \cdot (\hat{\Lambda}_2) \cdot \dots \cdot (\hat{\Lambda}_p)$$

Então:

$$(1,85)(0,15) = 0,28 \Rightarrow \text{determinante da matriz } R.$$

A proporção da variância total explicada por cada componente principal é dada por (1.19). Como sabido, os autovalores estão em ordem decrescente segundo seus valores, logo, a explicação de cada componente principal será respectivamente:

$$\frac{1,85}{2} \cdot 100 = 92,50\% \Rightarrow \hat{\Lambda}_1;$$

$$\frac{0,15}{2} \cdot 100 = 7,50\% \Rightarrow \hat{\Lambda}_2.$$

Então, tem-se que, a variância distribuída entre $\hat{\Lambda}_1 = 1,85$ e $\hat{\Lambda}_2 = 0,15$ está explicada 92,50% pelo primeiro eixo fatorial e 7,50% pelo segundo. Sendo assim, cada componente principal sintetiza a máxima proporção de variância contida nos dados da matriz de dados.

O cálculo da primeira componente principal referente à $\hat{\Lambda}_1 = 1,85$ será dado pelo autovetor associado a $\hat{\Lambda}_1$, sendo que a equação característica dos autovetores é $|R - \hat{\Lambda}_1 I| \hat{e}_1 = 0$.

Substituindo-se pelas matrizes R , I , pelo primeiro autovetor e pela matriz de incógnitas \hat{e}_i , têm-se:

$$\left[\begin{array}{cc} 1 & 0,85 \\ 0,85 & 1 \end{array} \right] - 1,85 \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] \begin{bmatrix} \hat{e}_{11} \\ \hat{e}_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Multiplicando-se o autovalor $\hat{\Lambda}_1$ pela matriz I e subtraindo da matriz R , chega-se às seguintes matrizes:

$$\left[\begin{array}{cc} 1-1,85 & 0,85 \\ 0,85 & 1-1,85 \end{array} \right] \begin{bmatrix} \hat{e}_{11} \\ \hat{e}_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Tais matrizes dão origem ao seguinte sistema:

$$\begin{cases} -0,85 \hat{e}_{11} + 0,85 \hat{e}_{12} = 0 \\ 0,85 \hat{e}_{11} - 0,85 \hat{e}_{12} = 0 \end{cases}.$$

Tal sistema de equações é indeterminado, em virtude de $|R - \hat{\Lambda} I| = 0$. Sendo assim, parte-se para:

$$\begin{bmatrix} -0,85 & 0,85 \\ 0,85 & -0,85 \end{bmatrix} = 0;$$

ou, equivalendo dizer que quando o vetor passa pela origem, fica $\hat{e}_{11} = \hat{e}_{12} = 0$.

Devido a isto, pode-se deixar uma das equações – a segunda –, e atribuir um valor arbitrário, não nulo, a uma das incógnitas – $\hat{e}_{12} = 1$. Dessa forma, tem-se que:

$$-0,85 \hat{e}_{11} + 0,85(1) = 0$$

$$-0,85 \hat{e}_{11} = -0,85, \text{ logo o valor da incógnita } \hat{e}_{11} \text{ será } 1.$$

O autovetor associado ao primeiro $\hat{\Lambda}_1 = 1,85$ será:

$$\hat{e}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ e sua norma será dada por:}$$

$$\|\hat{e}_1\| = \sqrt{(1)^2 + (1)^2} = 1,41.$$

Para que esse vetor seja unitário, é necessário normalizar o autovetor a 1, da seguinte maneira:

$$e_i = \frac{1}{\|\hat{e}_i\|} \cdot \hat{e}_i \quad (1.25)$$

Substituindo-se essa expressão pelos seus respectivos valores tem-se:

$$e_1 = \frac{1}{\|\hat{e}_1\|} \cdot \hat{e}_1 = e_1 = \frac{1}{\|1,41\|} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

logo, o primeiro autovetor normalizado será:

$$e_1 = \begin{bmatrix} 0,71 \\ 0,71 \end{bmatrix},$$

e a sua norma será:

$$\|e_1\| = \sqrt{(0,71)^2 + (0,71)^2} = 1$$

Assim sendo, a primeira componente principal equivale a: $Y_1 = 0,71Z_1 + 0,71Z_2$.

Resta encontrar a segunda componente principal, que é oriunda da raiz $\hat{\Lambda}_2 = 0,15$, dado que:

$$|R - \hat{\Lambda}_2 I| \hat{e}_2 = 0.$$

Novamente, devem-se substituir na equação as matrizes R , I , pelo segundo autovetor $\hat{\Lambda}_2 = 0,15$, seguido da matriz de incógnitas \hat{e}_i , obtendo-se:

$$\begin{bmatrix} 1 & 0,85 \\ 0,85 & 1 \end{bmatrix} - 0,15 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{e}_{21} \\ \hat{e}_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Multiplicando-se o autovalor $\hat{\Lambda}_2$ pela matriz I e subtraindo da matriz R , chegm-se às seguintes matrizes:

$$\begin{bmatrix} 1-0,15 & 0,85 \\ 0,85 & 1-0,15 \end{bmatrix} \begin{bmatrix} \hat{e}_{21} \\ \hat{e}_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Tais matrizes dão origem ao seguinte sistema:

$$\begin{cases} 0,85 \hat{e}_{21} + 0,85 \hat{e}_{22} = 0 \\ 0,85 \hat{e}_{21} + 0,85 \hat{e}_{22} = 0 \end{cases}.$$

Tal sistema de equações é indeterminado, em virtude de $|R - \hat{\Lambda}I| = 0$, sendo assim, parte-se para:

$$\begin{bmatrix} 0,85 & 0,85 \\ 0,85 & 0,85 \end{bmatrix} = 0.$$

Efetuada analogamente o procedimento anterior, tem-se:

$$0,85 \hat{e}_{21} + 0,85(1) = 0$$

$0,85 \hat{e}_{21} = -0,85$, logo o valor da incógnita \hat{e}_{21} será -1 .

O autovetor associado ao segundo $\hat{\Lambda}_2 = 0,15$ será:

$$\hat{e}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ e sua norma será dada por:}$$

$$\|\hat{e}_2\| = \sqrt{(-1)^2 + (1)^2} = 1,41.$$

Para que esse vetor seja unitário, é necessário normalizar o autovetor a 1, da seguinte maneira:

$$e_2 = \frac{1}{\|\hat{e}_2\|} \cdot \hat{e}_2.$$

Substituindo-se essa expressão pelos seus respectivos valores, tem-se:

$$x_2 = \frac{1}{\|1,41\|} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

logo, o segundo autovetor normalizado será:

$$e_2 = \begin{bmatrix} -0,71 \\ 0,71 \end{bmatrix},$$

e a sua norma será:

$$\|x_2\| = \sqrt{(-0,71)^2 + (0,71)^2} = 1$$

Assim sendo, a segunda componente principal equivale a: $Y_2 = -0,71Z_1 + 0,71Z_2$.

Sendo que os dois vetores são ortogonais, pois $\vec{e}_1 \vec{e}_2 = 0$ – o produto escalar é igual a zero –, equivalendo a uma segunda restrição feita por Hoffmann (1999, p. 5).

Para que seja satisfeita, deve-se multiplicar o primeiro autovetor normalizado transposto pelo segundo autovetor normalizado, ou seja:

$$e_1' e_2 = [0,71 \quad 0,71] \begin{bmatrix} -0,71 \\ 0,71 \end{bmatrix}.$$

Multiplicando-se os autovalores normalizados, tem-se a seguinte expressão:

$$\vec{e}_1 \vec{e}_2 = (0,71)(-0,71) + (0,71)(0,71) = 0$$

O passo seguinte é encontrar os escores procedendo-se de forma análoga feito com os resultados oriundos da matriz R , assim têm-se os resultados apresentados na Tabela 11.

Tabela 11 – Matriz de dados originais e a nova matriz gerada – MR.

Observações	Variáveis originais		Novas variáveis	
	X_1	X_2	Y_1	Y_2
1	62	61	87,33	-0,71
2	71	58	91,59	-9,23
3	69	94	115,73	17,75
4	55	35	63,90	-14,20
5	50	18	48,28	-22,72

Fonte: Elaboração própria.

Efetuada a análise de componentes principais, deve-se identificar ou mesmo, entender o significado de cada componente no estudo que está sendo realizado. A interpretação de cada componente principal é feita mediante seu grau de contribuição, ou seja, o grau de influência que cada variável tem sobre cada componente, sendo que esta importância é dada pela correlação entre cada variável padronizada Z_j e a componente Y_i .

Então, para finalizar a análise de componentes principais, é necessário fazer a correlação entre as variáveis Z_j e Y_i , como se pode verificar a seguir:

$$\begin{aligned}
 r_{x_1z_1} &= e_{11}\sqrt{\hat{\Lambda}_1} & r_{x_2z_1} &= e_{21}\sqrt{\hat{\Lambda}_2} \\
 r_{x_1z_1} &= 0,71\sqrt{1,85} = 0,97; & r_{y_2z_1} &= -0,71\sqrt{0,15} = -0,27; \\
 r_{x_1z_2} &= e_{12}\sqrt{\hat{\Lambda}_1} & r_{x_2z_2} &= r_{z_1y_2} = e_{22}\sqrt{\hat{\Lambda}_2} \\
 r_{x_1z_2} &= 0,71\sqrt{1,85} = 0,97; & r_{x_2z_2} &= 0,71\sqrt{0,15} = 0,27.
 \end{aligned}$$

Observa-se que a correlação da componente Y_1 é a mais significativa, corroborando, novamente, com as exigências de 1.18 e com o que foi concluído através da matriz S .

A Tabela 12 mostra os componentes principais encontrados no exemplo exploratório, os autovalores, autovetores, a correlação entre as variáveis, a percentagem de explicação de cada componente e a percentagem total da variância acumulada pelas componentes principais calculadas através da matriz de variância-covariância.

Tabela 12 – Resumo da análise de componentes principais pela matriz de correlação.

Componentes principais	Autovalores	Autovetores		Correlação entre Z_j e Y_i		Percentagem da variância de Y_i	Percentagem acumulada a variância dos Y_i
		Z_1	Z_2	Z_1	Z_2		
Y_1	1,85	0,71	0,71	0,97	0,97	92,50	92,50%
Y_2	0,15	-0,71	0,71	-0,27	0,27	7,50	100%

Fonte: Elaboração própria.

Ao findar o longo, porém indispensável exemplo de familiarização com a complexidade do método, passa-se para a análise fatorial e suas peculiaridades.

4.3 Análise Fatorial (AF)

Conforme Rummel (1967, p. 445), a Análise Fatorial “toma milhares e, potencialmente, milhões de medidas quantitativas e determina a ocorrência de padrões

distintos entre elas”, o que responde à questão qual: padrão de relação explica a relação entre as variáveis. Então, a Análise Fatorial objetiva descobrir fatores que expliquem o que existe de comum nas variáveis originais.

Reis (1997, p. 253) enfatiza que o objeto principal da AF é identificar novas variáveis – fatores –, em um número menor que o conjunto inicial, mas sem perda significativa de informações contidas no conjunto inicial das p variáveis, além de classificar e reduzir observações

[...] pode ser utilizada também para explorar qualquer área de conhecimento, mapear conceitos, estruturar domínios/predominâncias, separar e transformar dados, definir relações, testar alguma hipótese, mapear padrões sociais, formular teorias, controlar variáveis e, até mesmo, fazer inferências” (RUMMEL, 1967, p. 448-51).

A técnica, mais uma vez, tem origem em outra ciência que não a Economia. Spearman, um psicólogo, na busca de um fator que explicasse a inteligência, desenvolveu este instrumento para análise das faculdades humanas. Thurstone, outro psicólogo, foi além, por acreditar existir mais de um fator que explicasse a inteligência, “possibilitando o desenvolvimento da Análise Fatorial que permitia detectar mais do que um fator subjacente aos dados iniciais” (REIS, 1997, p. 253). Rummel (1967, p. 451) resgata o psicólogo Raymond Cattell e sua obra seminal sobre padrões de cultura obtidos por meio de Análise Fatorial em 1949.

A AF permite tornar os dados observados mais claros. Isto é possível pelo fato de serem descritas por um grupo de categorias básicas, em número menor que as variáveis originais, chamados fatores.

Os fatores são variáveis não observadas, escalas, itens, ou uma medida de qualquer espécie. Assim, estes fatores explicam a variância das variáveis originais que se revelam pelas correlações entre as variáveis que estão em análise.

O método de extração destes fatores, conforme Rummel (1967) é feito pela Análise das Componentes Principais, descrita na seção anterior, que se baseia no pressuposto de que pode definir \bar{X} vetores estatisticamente não correlacionados, a partir de combinações lineares dos p indicadores iniciais.

Assim, o exemplo da seção anterior mostrou que a ACP transforma um conjunto de variáveis correlacionadas entre si, num outro conjunto de variáveis não correlacionadas (ortogonais), que são as componentes principais, que resultam das combinações lineares do conjunto inicial.

Ambas, ACP e AF são técnicas de análise multivariada, aplicadas a um conjunto de variáveis, para descobrir quais dessas são mais relevantes na composição de cada fator, sendo estes independentes uns dos outros. Os fatores, que são gerados, serão utilizados de maneira representativa do processo em estudo no presente trabalho, e, para posteriores dados dos Censos a serem levantados no Rio Grande do Sul.

A ACP objetiva encontrar funções matemáticas, entre as variáveis iniciais, que expliquem o máximo possível de variação existente nos dados e permita descrever e reduzir estas. Já AF se baseia na explicação das covariâncias entre as variáveis. Assim, os fatores expressam o que existe de comum nas variáveis originais (RUMMEL, 1967, p. 455).

Sendo assim, é possível identificar fatores em um determinado conjunto de medidas e, desta forma, reduzir o número de um grande conjunto de variáveis em n fatores, onde os fatores, agrupam as variáveis conforme suas características e determinam a explicação de cada variável para os n fatores.

Para eficácia do método, além de se exigir uma correlação considerável entre as variáveis observadas, é necessário um número considerável de observações. Acima de tudo, o número de *cases* deve ser sempre superior ao de variáveis (ZAMBRANO & LIMA, 2004).

Desta forma, oportunamente Zambrano & Lima (2004, p. 556-7) salientam quatro etapas para realizar a técnica: a) normalização e verificação da adequação dos dados através da obtenção da matriz R , seguido dos testes KMO e Esfericidade de Barlett; b) determinação do número de fatores necessários para representar o conjunto; c) rotação dos fatores e d) cálculo dos escores fatoriais.

A primeira etapa, dos testes de adequação, está descrita na seção 4.2.1. Contudo, em virtude do elevado número de *cases* e variáveis, foi necessário o emprego do *software Statistical Package for Social Sciences* (SPSS) para realização do presente trabalho. Desta forma, recomenda-se o SPSS para efetuar estes testes, pois, o *Statistica* não os realiza.

A segunda etapa envolve a decisão do número de fatores necessários para representar o conjunto de dados. Como visto anteriormente, da matriz R são obtidos os autovalores e as proporções da variância total, utilizando-se o método de ACP. Desta forma, os fatores são combinações lineares de variáveis, sendo o primeiro o que está associado à maior proporção da variância dos dados originais; tantos componentes quantos forem necessários até que se tenha explicado uma proporção razoável da variância.

Então, geometricamente, os indicadores expressos em termos de vetores num sistema de coordenadas ortogonais agrupam-se conforme seu relacionamento. Admitindo-se que por

esses agrupamentos de vetores, passam eixos, os fatores na projeção dos vetores sobre os eixos, definem as cargas fatoriais – *factor loadings* – dos indicadores sobre si.

Conforme Hair Jr. (2005, p. 107) “uma carga fatorial representa uma correlação entre uma variável original e seu fator”. Essas seguem uma interpretação semelhante a determinação da significância estatística de coeficientes de correlação. Ainda, Hair Jr. (2005) salienta que as cargas fatoriais têm erros-padrão maiores do que as correlações normais; assim, essas devem ser avaliadas em níveis mais restritos. O autor sugere tamanhos de amostras necessárias para cada valor de carga fatorial, seja considerado significativo, conforme Tabela 13, sendo a significância por ele proposta baseada em um nível de significância (α) de 0,05, um nível de poder de 80% e erros-padrão, os quais, se pressupõe o dobro dos coeficientes de correlação convencionais.

Tabela 13 – Orientação para validação das cargas fatoriais com base no tamanho da amostra.

Carga fatorial	Tamanho necessário da amostra para significância
0,30	350
0,35	250
0,40	200
0,45	150
0,50	120
0,55	100
0,60	85
0,65	70
0,70	60
0,75	50

Fonte: Hair Jr. (2005, p. 107).

Assim, cada fator encontrado, explica uma parcela da variância dos dados originais, e, como estão ordenados de forma decrescente, o primeiro fator é o que vai explicar a maior proporção da variância dos dados, e assim sucessivamente como já demonstrado no exemplo da seção 4.2.1.

Assim, após identificar as variáveis que estão mais carregadas em um fator, entra o conhecimento do pesquisador sobre seu objeto, pois, os fatores serão por ele “batizados”. Assim, criam-se fatores que explicam o que existe de comum entre as variáveis com a menor perda de informação possível. De certa forma, esta pode ser a fase mais difícil do trabalho, pois, alguns fatores podem conter variáveis diversas com iguais comportamentos dificultando que se escolha um nome que contemple o poder de explicação do fator.

Ainda, determinam-se as comunalidades, isto é, a soma dos quadrados das cargas fatoriais dos indicadores sobre cada fator, indicando a eficiência destes na explicação da variabilidade total. Quanto mais próximo de 1 for o valor da comunalidade, maior a parte de variância da variável que estará sendo captada pelo conjunto de fatores considerados.

Resta ainda determinar o número de fatores a serem utilizados. Para o caso da utilização da matriz R , tem que ser observado o critério da raiz latente ou autovalores, pois, segundo Zambrano & Lima (2004, p. 559), “extraem-se apenas os fatores associados às características maiores que 1, já “os menores que um são descartados” (HAIR Jr., 2005, p. 101).

Na terceira etapa, já selecionados os fatores e batizados, os mesmos são submetidos a uma rotação, mantendo a ortogonalidade entre eles. “O objetivo da rotação ortogonal consiste em transformar os fatores obtidos em novos fatores independentes” (ZAMBRANO & LIMA, 2004, p. 559). A rotação altera a parcela de contribuição de cada fator, mas mantém os valores das comunalidades e a proporção de variância explicada pelo conjunto de fatores, pois, ela maximiza a carga de uma variável em um único fator, permitindo que, assim, melhor representem cada fator.

A Rotação *Varimax* foi o método escolhido para rotacionar os eixos por ser um método bastante utilizado e difundido na literatura consultada. Esse “forma um novo sistema de eixos ortogonais com o mesmo número de fatores e permite que o grupo de variáveis apareça com maior nitidez” (ZAMBRANO & LIMA, 2004, p. 560), facilitando a interpretação e análise. O método maximiza a soma das variâncias das cargas fatoriais e busca a simplificação das colunas da matriz fatorial. A simplificação máxima é conseguida se houver apenas cargas fatoriais próximas de 1 ou 0. Quando as correlações são próximas de +1 ou -1 indicam clara associação positiva ou negativa e, próximas de zero, indicam falta de associação entre fator e variável.

E, por fim, são calculados os escores fatoriais para cada observação. O escore fatorial, ou índice, é obtido pela multiplicação do valor padronizado da variável X_j pelo coeficiente do escore fatorial correspondente. Esses valores constituem, exatamente, as projeções normalizadas dos pontos representativos das unidades sobre os eixos dos fatores que, por construção, tem média zero e desvio-padrão unitário. Dessa forma, Hair Jr. (2005) salienta que elevados escores fatoriais positivos ou negativos mostrarão que tal observação tem alta influência daquele fator.

5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Esta seção se dedica a discutir e analisar os resultados advindos da aplicação das técnicas anteriormente expostas. Dessa forma, em um primeiro momento, agrupam-se os municípios dos COREDEs agrícolas do Rio Grande do Sul (RS), para, em seguida, explicar a peculiaridade dos grupos formados. A explicação será efetuada na segunda subseção que se discutem, já agrupados, os indivíduos, os níveis de modernização, segundo os fatores encontrados.

5.1 Análise de Agrupamento (AA)

Da forma como foi feita a revisão da literatura sobre análise multivariada, os resultados também são apresentados inicialmente pelo agrupamento das observações. Pois, primeiro serão agrupados e, posteriormente, explicado a relação desses grupos homogêneos almejados para o estudo na seção 5.2.

Primando a Análise de Agrupamentos, o primeiro dendrograma obtido provou a teoria de que no maior salto haveria uma considerável perda de informações, formando três grupos, sendo o corte efetuado entre as distâncias 275,06 e 675,59, anterior à fusão total – 1664,40. Um grupo I foi formado com 5,76% da amostra, ou seja, oito municípios, sendo sete do COREDE Fronteira Oeste⁴¹ – Alegrete, Itaqui, Rosário do Sul, Santana do Livramento, São Borja, São Gabriel, Uruguaiana – mais um; Cachoeira do Sul do COREDE Jacuí Centro⁴². O grupo II continha 10,79% e o terceiro grupo, o maior, com 83,45% da amostra, composto por 116 municípios.

O grupo II formou-se com um total de 15 municípios: Giruá do COREDE Missões⁴³; Ajuricaba, Augusto Pestana, Ijuí, Panambi, Crissiumal e Três Passos do COREDE Noroeste

⁴¹ Alegrete, Itacurubi, Itaqui, Manoel Viana, Quaraí, Rosário do Sul, Santana do Livramento, São Borja, São Gabriel e Uruguaiana.

⁴² Cachoeira do Sul, Cerro Branco, Paraíso do Sul, Restinga Seca, São Sepé e Vila Nova do Sul.

⁴³ Bossoroca, Caibaté, Cerro Largo, Dezesseis de Novembro, Ente-Ijuís, Eugênio de Castro, Garruchos, Giruá, Guarani das Missões, Pirapó, Porto Xavier, Roque Gonzáles, Salvador das Missões, Santo Ângelo, Santo Antônio das Missões, São Luiz Gonzaga, São Miguel das Missões, São Nicolau, São Paulo das Missões, São Pedro do Butiá e Vitória das Missões.

Colonial⁴⁴; Sananduva e Tapejara do COREDE Nordeste⁴⁵; Cruz Alta e Ibirubá do COREDE Alto Jacuí⁴⁶; os municípios de Barros Cassal, Espumoso, Fontoura Xavier e Soledade do COREDE Alto da Serra do Botucaraí⁴⁷.

Já o grupo III, formou-se do COREDE Médio Alto Uruguai⁴⁸ na totalidade e o restante dos municípios dos COREDEs em questão. *A priori*, parece fazer sentido dar o corte no ponto “recomendado pela teoria”, contudo, o grupo III não é tão homogêneo quanto se esperava além do elevado tamanho, que ocasiona uma perda vital de informações.

Isto coaduna com a idéia de que no maior salto, dado corte, se dá a maior perda de informações. A fusão começa com o município de Bom Progresso e Inhacorá do COREDE Noroeste Colonial na distância 0,35 e, o último grupo se funde na distância 1664,40. A Figura 6 traz o gráfico das distâncias de fusão dos indivíduos em grupos, a elipse em vermelho delimita o maior salto no processo de agrupamento.

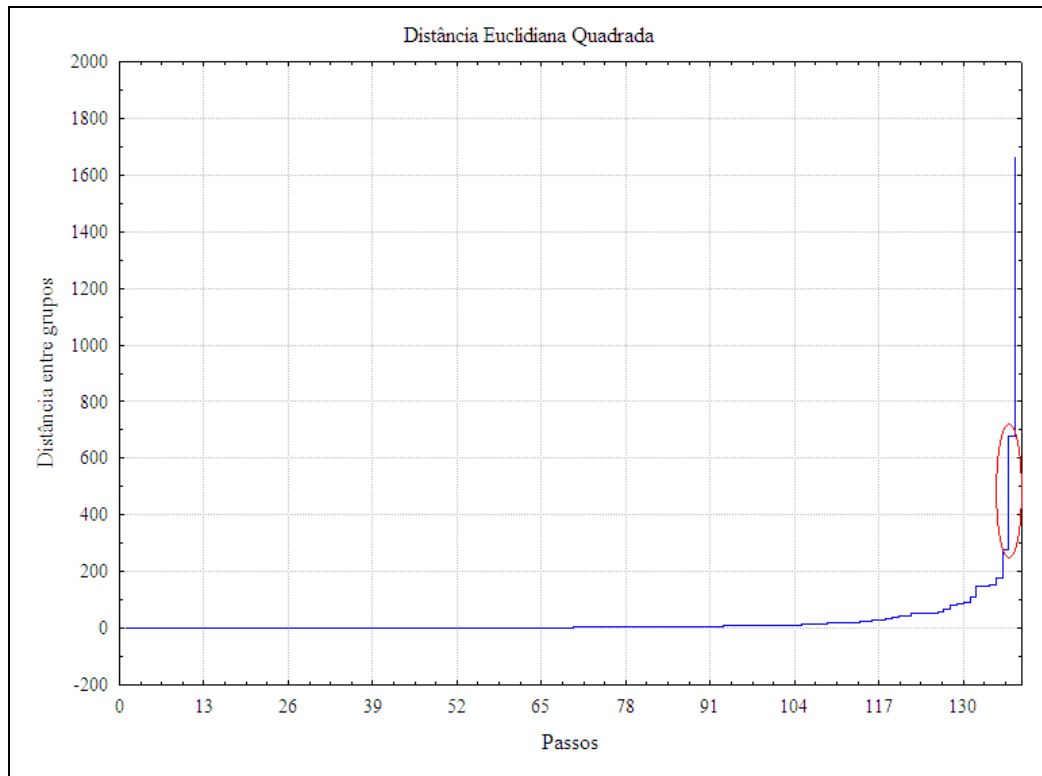
⁴⁴ Ajuricaba, Augusto Pestana, Barra do Guarita, Bom Progresso, Braga, Campo Novo, Catuípe, Chiapeta, Condor, Coronel Barros, Coronel Bicaco, Crissiumal, Derrubadas, Humaitá, Ijuí, Inhacorá, Jóia, Miraguaí, Panambi, Pejuçara, Redentora, Santo Augusto, São Martinho, São Valério do Sul, Sede Nova, Tenente Portela, Tiradentes do Sul, Três Passos e Vista Gaúcha.

⁴⁵ Água Santa, André da Rocha, Barracão, Cacique Doble, Caseiros, Esmeralda, Ibiaçá, Ibiraiaras, Lagoa Vermelha, Machadinho, Maximilhano de Almeida, Paim Filho, Sananduva, Santo Expedito do Sul, São João da Urtiga, São José do Ouro, Tapejara e Tupanci do Sul.

⁴⁶ Colorado, Cruz Alta, Fortaleza dos Valos, Ibirubá, Lagoa dos Três Cantos, Não-Me-Toque, Quinze de Novembro, Saldanha Marinho, Salto do Jacuí, Santa Bárbara do Sul, Selbach e Tapera.

⁴⁷ Alto Alegre, Barros Cassal, Campos Borges, Espumoso, Fontoura Xavier, Gramado Xavier, Ibirapuitã, Itapuca, Lagoão, Mormaço, Nicolau Vergueiro, São José do Herval, Soledade e Victor Graeff.

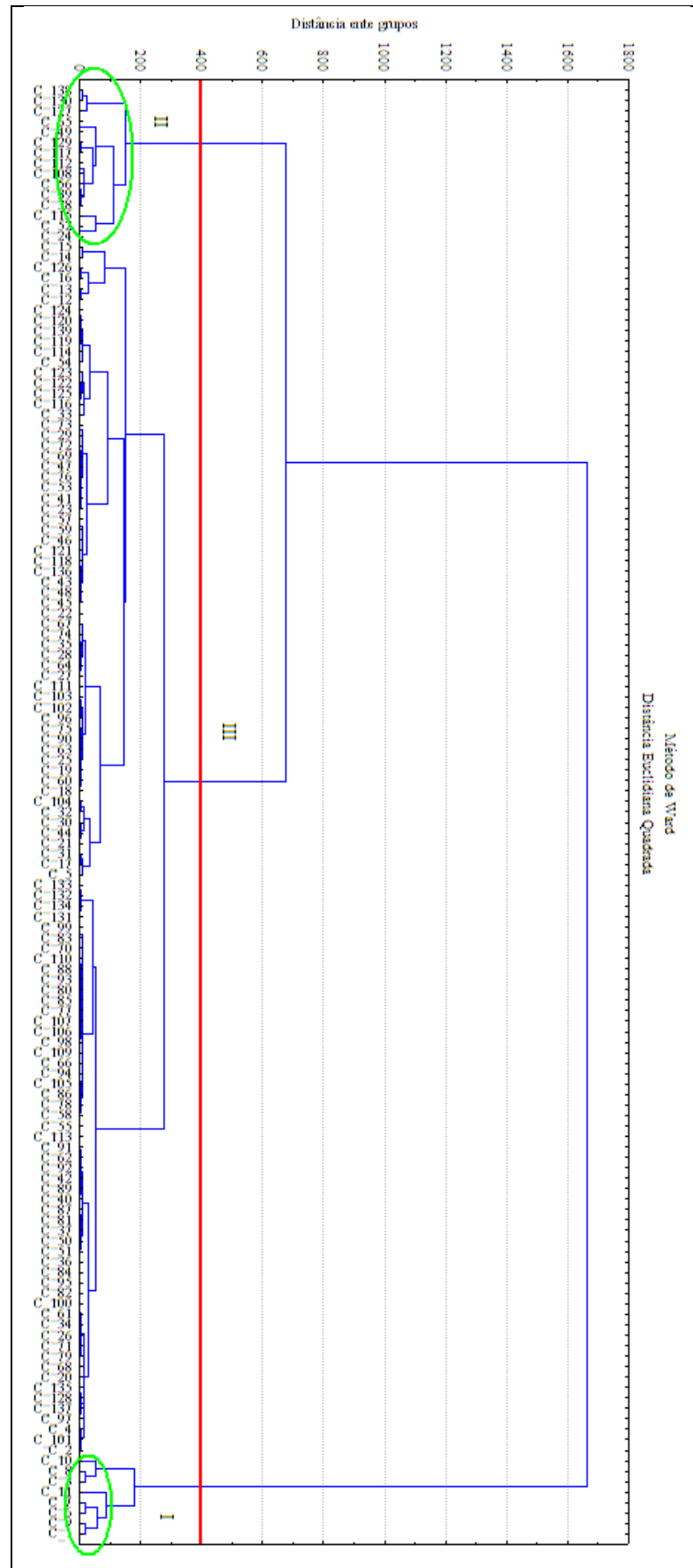
⁴⁸ Alpestre, Ametista do Sul, Boa Vista das Missões, Caiçara, Cerro Grande, Dois Irmãos das Missões, Engenho Velho, Erval Seco, Frederico Westphalen, Gramado dos Loureiros, Iraí, Jaboticaba, Lajeado do Bugre, Liberato Salzano, Nonoai, Novo Tiradentes, Palmitinho, Pinhal, Pinheirinho do Vale, Planalto, Rio dos Índios, Rodeio Bonito, Sagrada Família, Seberí, Taquaruçu do Sul, Três Palmeiras, Trindade do Sul, Vicente Dutra e Vista Alegre.



Fonte: Figura/output fornecido pelo software Statistica 7.0.

Figura 6 - Gráfico das distâncias de fusão dos grupos – maior salto.

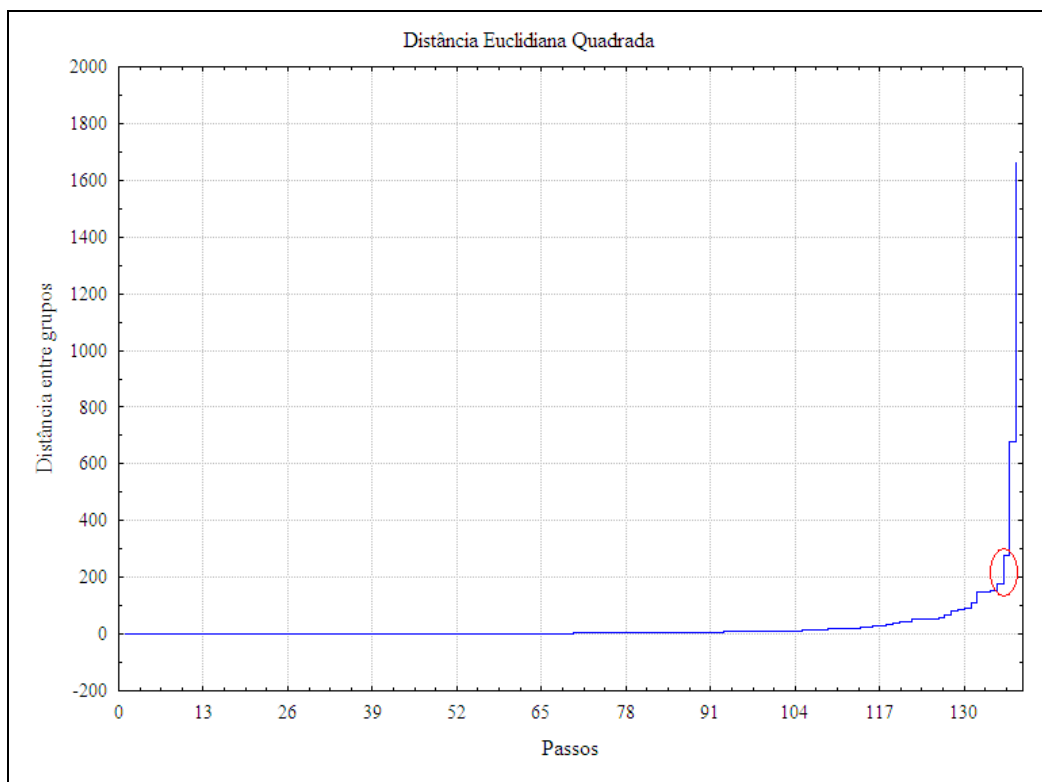
Observe na Figura 7 a Linha de Fenon – em vermelho – que delimita o corte no maior salto e, que, conseqüentemente deu origem aos grupos supracitados. Dessa forma, percebe-se que se formam três grupos – detalhe em verde – que somados contêm a totalidade da amostra. Contudo, o grupo III possui um tamanho demasiadamente elevado e, com isso, se perde uma quantia de informações cruciais para o estudo.



Fonte: Figura/output fornecido pelo *software* Statistica 7.0.

Figura 7 – Dendrograma – maior salto.

Prosseguindo na análise, ignorando o corte no maior salto, efetuou-se um novo corte entre 179,37 e 275,06. Este é o ponto que une dois grandes grupos que anteriormente originou o grupo III. Portanto, o corte sendo feito entre essas distâncias aumenta a homogeneidade entre os novos grupos formados, advindos do antigo grupo III. Então, têm-se: os antigos grupos I e II mantendo sua formação e forte homogeneidade de seus elementos mais dois novos grupos oriundos da divisão do grupo III. O corte foi efetuado conforme as distâncias salientadas e, fica mais claro na Figura 8:



Fonte: Figura/output fornecido pelo software Statistica 7.0.

Figura 8 - Gráfico das distâncias de fusão dos grupos.

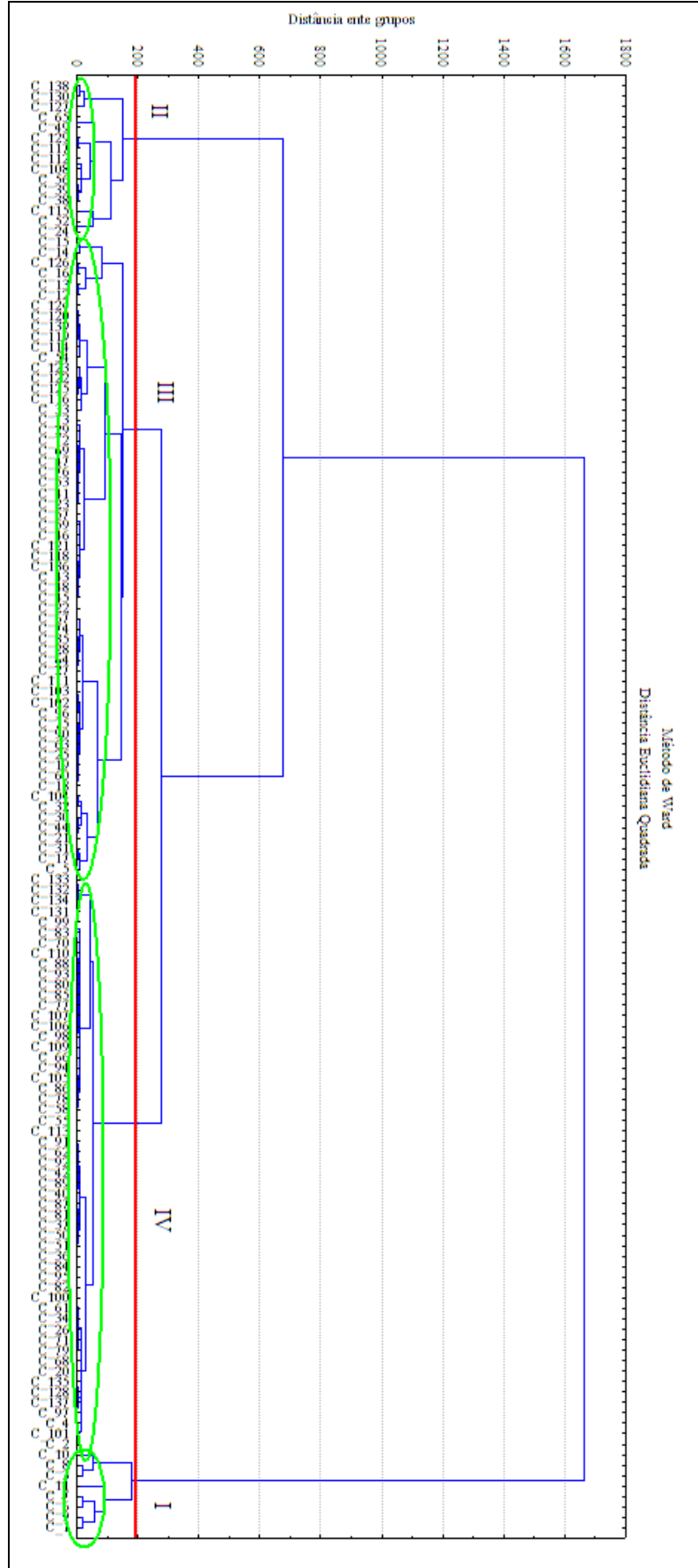
Para os novos grupos formados tem-se que o grupo III formou-se dos municípios de Quaraí, Cerro Branco, Paraíso do Sul, Restinga Seca, São Sepé, Vila Nova do Sul, Bossoroca, Caibaté, Cerro Largo, Entre-Ijuís, Eugênio de Castro, Garruchos, Guarani das Missões, Porto Xavier, Roque Gonzales, Salvador das Missões, São Paulo das Missões, Santo Ângelo, Santo Antônio das Missões, São Luiz Gonzaga, São Miguel das Missões, Bom Progresso, Campo Novo, Catuípe, Chiapeta, Condor, Coronel Barros, Coronel Bicaco, Inhacorá, Jóia, Pejuçara, Santo Augusto, São Martinho, Tenente Portela, Tiradentes do Sul, Alpestre, Boa Vista das

Missões, Dois Irmãos das Missões, Engenho Velho, Erval Seco, Frederico Westphalen, Gramado dos Loureiros, Seberi, Água Santa, Ibiaçá, Ibiraiaras, Lagoa Vermelha, São José do Ouro, Colorado, Fortaleza dos Valos, Lagoa dos Três Cantos, Não-Me-Toque, Quinze de Novembro, Saldanha Marinho, Salto do Jacuí, Santa Bárbara do Sul, Selbach, Tapera, Alto Alegre, Nicolau Vergueiro e Victor Graeff.

Por fim, o grupo IV tem como membros os municípios de Itacurubi, Manoel Viana, Dezesesseis de Novembro, Pirapó, São Nicolau, São Pedro do Butiá, Vitória das Missões, Barra do Guarita, Braga, Derrubadas, Humaitá, Miraguai, Redentora, São Valério do Sul, Sede Nova, Vista Gaúcha, Ametista do Sul, Caiçara, Cerro Grande, Iraí, Jaboticaba, Lajeado do Bugre, Liberato Salzano, Nonoai, Novo Tiradentes, Palmitinho, Pinhal, Pinheirinho do Vale, Planalto, Rio dos Índios, Rodeio Bonito, Sagrada Família, Taquaruçu do Sul, Três Palmeiras, Trindade do Sul, Vicente Dutra, Vista Alegre, André da Rocha, Barracão, Cacique Doble, Caseiros, Esmeralda, Machadinho, Maximiliano de Almeida, Paim Filho, Santo Expedito do Sul, São João da Urtiga, Tupanci do Sul, Campos Borges, Gramado Xavier, Ibirapuitã, Itapuca, Lagoão, Mormaço e São José do Herval.

Além do grupo I formado por: Alegrete, Itaqui, Rosário do Sul, Santana do Livramento, São Borja, São Gabriel, Uruguaiana e Cachoeira do Sul e, o grupo II, composto por: Giruá, Ajuricaba, Augusto Pestana, Ijuí, Panambi, Crissiumal, Três Passos, Sananduva, Tapejara, Cruz Alta, Ibirubá, Barros Cassal, Espumoso, Fontoura Xavier e Soledade.

Sendo assim, têm-se quatro grupos compreendem na totalidade os 139 municípios propostos para o estudo. Finalmente, encerrando esta primeira etapa do agrupamento, formaram-se o grupo I com oito municípios, o grupo II com 15 e o grupo III e IV, 61 e 55 municípios respectivamente. Grupos estes sugeridos pelo *software Statistica*, seguem o mesmo padrão da Figura 7, exceto, pela transformação do antigo grupo III em dois novos: os atuais grupos III e IV.



Fonte: Figura/output fornecido pelo software Statistica 7.0.

Figura 9 - Dendrograma.

5.2 Análise Fatorial (AF)

Cabe lembrar, que foram analisados 139 municípios do Estado do Rio Grande do Sul sob a ótica de 17 variáveis – discriminadas no Quadro 1 – sendo estes municípios pertencentes aos COREDEs essencialmente agrícolas⁴⁹, conforme classificação da Fundação de Economia e Estatística (FEE) junto ao Governo do Estado do Rio Grande do Sul – Rumos 2015 (2005). Dessa forma seguem-se os passos descritos por Zambrano & Lima (2004).

A primeira etapa consiste em verificar a adequação dos dados ao método de Análise Fatorial. Para isso, gerou-se a matriz $R_{17 \times 17}$ de correlações entre as variáveis. A Tabela 10 apresenta os valores obtidos.

Tabela 10 – Matriz de correlação entre as variáveis originais normalizadas.

Variáveis	POP	PEA	ASSPEC	ASSAGR	TRATOR	ADUBO	IRRIGA	RECVEG	RECANI	FINANC	MAQCOL	VALVEG	VALANI	AREXPL	RENDA	DISTRG	GADO
POP	1,00																
PEA	0,38	1,00															
ASSPEC	0,38	0,59	1,00														
ASSAGR	0,64	0,54	0,64	1,00													
TRATOR	0,27	0,64	0,57	0,52	1,00												
ADUBO	0,38	0,90	0,66	0,66	0,54	1,00											
IRRIGA	0,08	0,52	0,23	0,25	0,72	0,32	1,00										
RECVEG	0,25	0,51	0,39	0,36	0,86	0,29	0,67	1,00									
RECANI	0,14	0,59	0,47	0,26	0,72	0,35	0,66	0,75	1,00								
FINANC	0,12	0,14	0,23	0,17	0,62	-0,04	0,48	0,77	0,52	1,00							
MAQCOL	0,30	0,56	0,55	0,51	0,96	0,47	0,56	0,85	0,61	0,65	1,00						
VALVEG	0,26	0,57	0,43	0,38	0,87	0,36	0,70	0,99	0,76	0,74	0,86	1,00					
VALANI	0,16	0,60	0,40	0,21	0,75	0,32	0,72	0,78	0,96	0,56	0,64	0,80	1,00				
AREXPL	0,10	0,52	0,32	0,13	0,70	0,21	0,70	0,78	0,93	0,58	0,61	0,77	0,98	1,00			
RENDA	-0,09	0,01	0,31	0,13	0,43	0,08	0,14	0,29	0,19	0,36	0,47	0,29	0,19	0,16	1,00		
DISTRG	-0,25	-0,12	-0,14	-0,25	-0,34	-0,09	-0,43	-0,18	-0,24	-0,17	-0,28	-0,20	-0,31	-0,31	-0,14	1,00	
GADO	0,06	0,50	0,28	0,08	0,65	0,19	0,70	0,72	0,93	0,53	0,53	0,72	0,97	0,99	0,13	-0,30	1,00

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

De pronto, percebe-se um elevado grau de correlação entre a maioria das variáveis o que, visualmente, indicaria uma adequação dos dados. De forma a se verificar e ratificar esta suposição efetuou-se os testes *Kaiser-Meyer-Olkin Measure of Adequacy* (KMO) e de *Barlett of Sphericity* (BASP). O teste de KMO revelou um valor meritório de 0,816, indicando uma

⁴⁹ Sendo eles: a) Fronteira Oeste, com 10 municípios; b) Jacuí Centro, com 6 municípios; c) Missões, 21 municípios; d) Noroeste Colonial, 29 municípios; e) Médio Alto Uruguai, também, com 29 municípios, f) Noroeste, 18 municípios, g) Alto Jacuí e h) Alto da Serra do Botucará com, 12 e 14 municípios respectivamente.

boa adequação dos dados ao método. Por fim, o teste BASPH (4097,40) mostrou-se significativo a 1% ($p < 0,01$), assim, rejeitando-se a hipótese nula – $H_0 : P = 1$. Desse modo, ambos os testes realizados permitem concluir que a amostra é adequada ao procedimento de análise, ou seja, ao emprego da Análise Fatorial.

Prosseguindo na Análise dos Componentes Principais (ACP) a serem encontrados através da matriz R , num máximo de 17 componentes, a Tabela 12 apresenta os CP correlacionados com as variáveis originais normalizadas. Os valores em destaque representam a contribuição de cada variável em cada fator, ou seja, no fator 1, 0,323; 0,713; 0,589; 0,492; 0,934; 0,523; 0,768; 0,904; 0,879; 0,660; 0,867; 0,922; 0,903; 0,866 e 0,827. E, dessa forma, procede-se a análise analogamente para todos os fatores. Resta ainda delimitar o número de fatores necessários para análise.

As correlações já permitem tal delimitação, pois, lembrando que os fatores são apresentados em ordem decrescente, à medida que aumenta o número de fatores as correlações são menos significativas. Dessa forma, o primeiro é o fator que apresenta o maior número de correlações significativas com as variáveis, ou seja, que melhor explica essas. À medida que se avança, o fator 2 apresenta também correlações significativas, porém, em um número menor que o fator 1. Os fatores 3 e 4 apresentam apenas uma correlação significativa. Agora, se observado, do fator 5 ao fator 17 os valores das correlações tendem a diminuir, podendo ser aproximadamente nulas em alguns casos.

Em um primeiro olhar, a Tabela 15 fornece uma idéia de que, o número de fatores seria quatro para explicar o conjunto da variância das variáveis. Mas, para escolher o número de fatores existe outro critério, o de observar os autovalores maiores que 1, que será visto a seguir. Ainda, estão destacadas as correlações significativas superiores a 0,50, conforme Hair Jr. (2005), que estão dispostas na Tabela 13.

Tabela 15 – Matriz de correlação entre as variáveis originais normalizadas e as componentes principais – significativas > 0,50.

	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6	Fator 7	Fator 8	Fator 9	Fator 10	Fator 11	Fator 12	Fator 13	Fator 14	Fator 15	Fator 16	Fator 17
POP	0,323	-0,590	-0,024	-0,445	-0,424	0,247	-0,280	0,061	0,100	-0,115	0,044	-0,020	0,006	0,005	0,002	0,000	0,001
PEA	0,713	-0,459	-0,380	0,217	0,069	-0,148	-0,122	-0,053	0,152	0,080	-0,041	0,059	-0,099	-0,014	0,009	-0,002	-0,002
ASSPEC	0,589	-0,546	0,101	0,187	0,169	0,352	0,325	-0,062	0,083	-0,203	-0,014	0,028	-0,003	0,002	0,000	0,000	0,000
ASSAGR	0,492	-0,732	0,122	-0,179	-0,111	0,022	0,155	0,221	-0,238	0,172	-0,019	0,047	-0,009	-0,006	-0,003	0,000	0,000
TRATOR	0,934	-0,087	0,209	0,022	0,070	-0,167	-0,013	-0,079	-0,086	-0,037	0,147	-0,031	0,004	-0,015	0,071	-0,016	0,005
ADUBO	0,523	-0,717	-0,226	0,244	0,185	-0,172	-0,052	-0,028	0,103	0,104	-0,015	-0,057	0,096	0,026	-0,009	0,001	0,001
IRRIGA	0,768	0,196	-0,154	-0,175	0,149	-0,399	0,075	0,300	0,070	-0,183	0,045	0,000	-0,004	0,006	-0,018	0,002	-0,002
RECVEG	0,904	0,164	0,172	0,047	-0,270	-0,108	-0,011	-0,061	-0,066	-0,051	-0,159	-0,004	0,004	0,049	0,018	-0,003	-0,027
RECANI	0,879	0,216	-0,253	0,093	0,049	0,236	0,029	0,059	-0,071	0,049	0,006	-0,185	-0,049	0,012	-0,011	0,003	0,002
FINANC	0,660	0,343	0,438	-0,058	-0,313	0,012	0,191	0,042	0,280	0,183	0,052	-0,006	0,002	-0,002	0,000	-0,001	0,000
MAQCOL	0,867	-0,112	0,361	0,048	-0,035	-0,135	-0,064	-0,213	-0,104	-0,020	0,140	0,018	-0,018	0,011	-0,056	0,019	-0,002
VALVEG	0,922	0,115	0,138	0,062	-0,228	-0,136	-0,016	-0,056	-0,035	-0,067	-0,174	-0,011	0,015	-0,045	-0,008	0,004	0,027
VALANI	0,903	0,276	-0,248	0,016	0,055	0,167	-0,053	0,006	-0,019	0,022	0,019	0,021	0,043	-0,087	-0,011	-0,003	-0,021
AREXPL	0,866	0,381	-0,247	-0,007	0,024	0,164	-0,045	-0,012	-0,036	0,030	0,024	0,077	0,013	0,041	-0,020	-0,049	0,011
RENDA	0,322	0,047	0,724	0,163	0,464	0,173	-0,258	0,164	0,036	0,015	-0,055	0,010	-0,007	0,002	0,000	0,000	0,000
DISTRG	-0,351	0,005	-0,007	0,805	-0,429	0,033	-0,061	0,171	-0,038	-0,056	0,073	0,015	0,005	-0,001	-0,001	0,000	0,000
GADO	0,827	0,410	-0,313	-0,010	0,052	0,178	-0,035	0,032	-0,019	0,034	0,022	0,078	0,021	0,032	0,028	0,049	0,008

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

Dessa forma, a Tabela 16, corrobora com a compreensão dos dados contidos na Tabela 15, tendo por base o pressuposto de quatro possíveis fatores e suas correlações com os municípios propostos para o estudo.

Tabela 16 – Escores fatoriais para os grupos de I a IV*.

Grupos	Município	Fator 1	Fator 2	Fator 3	Fator 4
Grupo I	Alegrete	4,216	1,887	-2,614	1,223
	Itaqui	3,132	2,680	1,801	1,455
	Rosário do Sul	1,889	1,398	-1,756	-0,778
	Santana do Livramento	2,357	2,157	-3,763	-0,461
	São Borja	2,283	1,222	1,943	0,379
	São Gabriel	3,332	2,094	-1,770	-1,281
	Uruguaiana	4,464	2,609	0,780	1,097
	Cachoeira do Sul	3,551	-0,934	-1,587	-1,425
Grupo II	Giruá	1,673	-2,582	0,550	1,254
	Ajuricaba	0,689	-1,255	0,665	0,374
	Augusto Pestana	0,420	-1,236	0,470	0,345
	Crissiumal	0,618	-2,695	-1,052	2,112
	Ijuí	1,607	-2,435	0,841	0,762
	Panamby	0,574	-1,225	1,115	0,304
	Três Passos	0,646	-2,691	-1,218	2,254
	Sananduva	0,487	-1,329	-0,272	1,520
	Tapejara	0,517	-1,328	0,094	1,291
	Cruz Alta	2,835	-2,751	2,155	-0,783
	Ibirubá	1,258	-2,459	1,517	-0,766
	Barros Cassal	0,164	-2,915	-1,473	-2,617
	Espumoso	0,761	-2,428	0,458	-0,899
	Fontoura Xavier	0,091	-2,519	-1,523	-2,690
	Soledade	0,295	-1,132	-0,755	-1,439
Grupo III	Quaraí	0,481	1,680	-0,932	-0,522
	Cerro Branco	-0,254	-0,531	-0,388	-2,000
	Paraíso do Sul	-0,009	-0,476	-0,336	-2,365
	Restinga Seca	0,878	-0,113	0,126	-2,497
	São Sepé	1,575	0,075	-0,081	-2,115
	Vila Nova do Sul	-0,627	1,023	0,030	-2,508
	Bossoroca	0,255	0,866	0,351	-0,212
	Caibaté	-0,112	-0,500	-0,218	0,573
	Cerro Largo	-0,187	-0,596	0,310	0,692
	Entre-Ijuís	0,108	-0,072	0,611	0,203
	Eugênio de Castro	-0,345	0,371	0,591	-0,062
	Garruchos	-0,445	1,042	0,774	0,844
	Guarani das Missões	-0,027	-0,930	-0,347	0,529
	Porto Xavier	-0,472	-0,143	-0,536	1,026
	Roque Gonzáles	-0,234	-0,287	-0,533	0,708
	Salvador das Missões	-0,458	0,009	1,172	0,500
	Santo Ângelo	0,327	-0,871	0,119	0,349
	Santo Antônio das Missões	0,332	0,993	-0,734	0,146
	São Luiz Gonzaga	1,022	-0,384	-0,207	0,906
	São Miguel das Missões	0,617	0,898	1,656	0,118
	São Paulo das Missões	-0,327	-0,417	-0,887	0,809
	Bom Progresso	-0,754	0,673	0,731	0,251
	Campo Novo	-0,471	0,701	0,935	0,038
	Catuípe	0,212	0,046	0,424	0,053
	Chiapeta	-0,261	0,579	1,127	0,049
	Condor	0,057	0,206	1,398	-0,164
	Coronel Barros	-0,456	0,468	1,142	-0,230
	Coronel Bicaco	-0,088	0,084	0,727	0,309
	Inhacorá	-0,719	0,866	0,975	0,081
	Jóia	0,456	-0,235	1,059	0,002

Tabela 16 – Escores fatoriais para os grupos de I a IV*.

(Continua)

Grupos	Município	Fator 1	Fator 2	Fator 3	Fator 4
Grupo III	Pejuçara	0,024	0,536	1,860	-0,463
	Santo Augusto	-0,047	0,356	1,324	0,106
	São Martinho	-0,266	-0,304	0,026	0,576
	Tenente Portela	-0,134	-0,691	-0,425	0,825
	Tiradentes do Sul	-0,236	-0,564	-0,814	1,214
	Alpestre	-0,245	-1,139	-1,685	1,025
	Boa Vista das Missões	-0,560	0,886	1,626	0,093
	Dois Irmãos das Missões	-0,490	0,931	1,513	0,392
	Engenho Velho	-0,597	0,287	1,064	0,402
	Erval Seco	-0,353	-0,292	-0,928	0,363
	Frederico Westphalen	-0,085	-0,664	-0,044	0,546
	Gramado dos Loureiros	-0,760	0,907	0,925	0,601
	Seberi	-0,172	-0,654	-0,533	0,303
	Água Santa	-0,142	-0,061	-0,223	0,670
	Ibiaçá	-0,075	-0,316	0,335	0,751
	Ibiraiaras	-0,055	-0,488	0,257	0,841
	Lagoa Vermelha	0,663	-0,304	-0,861	1,218
	São José do Ouro	-0,210	-0,252	0,460	1,007
	Colorado	-0,020	-0,326	1,512	-0,822
	Fortaleza dos Valos	0,253	0,253	1,771	-0,734
	Lagoa dos Três Cantos	-0,436	0,335	1,109	-0,550
	Não-Me-Toque	0,314	-0,605	1,508	-0,484
	Quinze de Novembro	-0,122	-0,363	0,545	-0,691
	Saldanha Marinho	-0,326	0,321	1,283	-0,564
	Salto do Jacuí	0,009	0,298	0,760	-1,291
	Santa Bárbara do Sul	0,732	0,293	2,334	-0,944
	Selbach	-0,047	-0,541	1,019	-0,812
	Tapera	-0,016	0,186	1,779	-0,816
	Alto Alegre	-0,424	0,201	0,120	-4,035
	Nicolau Vergueiro	-0,496	0,632	1,033	-0,378
	Victor Graeff	0,086	-0,731	1,516	-0,280
Grupo IV	Itacurubi	-0,275	1,193	-0,749	-0,483
	Manoel Viana	-0,162	1,142	0,194	-0,202
	Dezesseis de Novembro	-0,835	0,591	-0,916	0,179
	Pirapó	-0,718	0,468	-0,521	0,494
	São Nicolau	-0,530	0,676	-0,408	0,304
	São Pedro do Butiá	-0,488	0,176	0,263	0,568
	Vitória das Missões	-0,530	0,307	-0,072	0,043
	Barra do Guarita	-0,906	0,831	0,184	0,562
	Braga	-0,692	0,538	0,089	0,055
	Derrubadas	-0,449	-0,103	0,055	0,875
	Humaitá	-0,435	-0,060	0,284	0,543
	Miraguaí	-0,703	0,169	-1,047	0,347
	Redentora	-0,492	-0,044	-0,766	0,062
	São Valério	-0,824	0,650	-0,454	-0,236
	Sede Nova	-0,584	0,273	0,322	0,377
	Vista Gaúcha	-0,592	0,022	-0,435	0,540
	Ametista do Sul	-0,957	0,772	-0,789	-0,281
	Caiçara	-0,442	-0,511	-0,636	0,028
	Cerro Grande	-0,831	0,474	-0,862	-0,157
	Iraí	-0,549	-0,060	-0,372	0,507
	Jaboticaba	-0,684	0,392	-0,688	-0,308
	Lajeado do Bugre	-0,964	0,863	-1,022	-0,575
	Liberato Salzano	-0,526	-0,281	-0,922	0,497
	Nonoai	-0,432	0,386	0,303	0,715
	Novo Tiradentes	-0,842	0,535	-0,635	-0,253
	Palmitinho	-0,465	-0,446	-0,826	0,306
	Pinhal	-0,823	0,629	-0,167	-0,193
	Pinheirinho do Vale	-0,580	-0,009	-0,518	0,249
	Planalto	-0,569	-0,025	-0,979	0,670
	Rio dos Índios	-0,525	0,112	0,081	0,832
	Rodeio Bonito	-0,526	-0,217	-0,488	0,285
Sagrada Família	-0,858	0,746	-0,238	0,004	

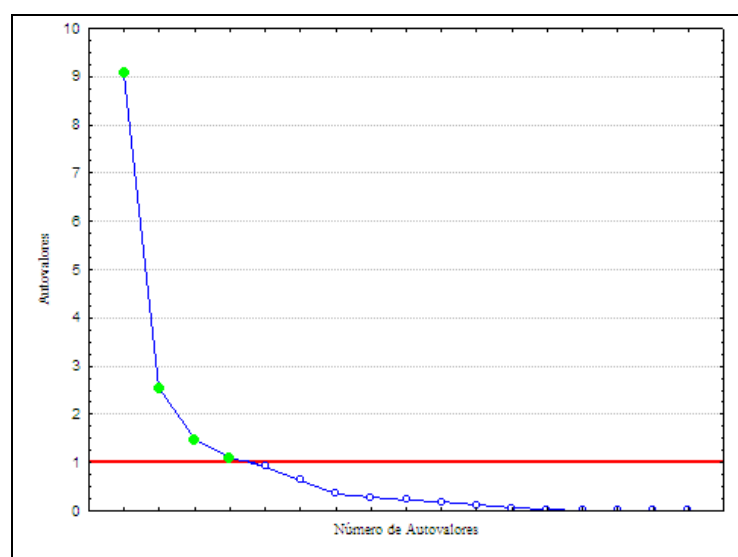
Tabela 16 – Escores fatoriais para os grupos de I a IV*. (Continua)

Grupos	Município	Fator 1	Fator 2	Fator 3	Fator 4
Grupo IV	Taquaruçu do Sul	-0,684	0,500	0,302	0,086
	Três Palmeiras	-0,603	0,461	0,082	0,410
	Trindade do Sul	-0,514	0,138	-0,473	0,618
	Vicente Dutra	-0,602	-0,130	-1,273	0,150
	Vista Alegre	-0,738	0,375	-0,371	-0,001
	André da Rocha	-0,742	1,145	-0,218	-0,811
	Barracão	-0,436	0,249	-0,238	0,948
	Cacique Doble	-0,495	-0,526	-0,894	0,714
	Caseiros	-0,723	0,713	-0,365	0,176
	Esmeralda	-0,380	0,698	-1,346	0,173
	Machadinho	-0,543	0,191	-1,192	1,007
	Maximiliano de Almeida	-0,554	0,107	-0,588	1,105
	Paim Filho	-0,526	0,167	-0,320	1,014
	Santo Expedito do Sul	-0,688	0,137	-0,355	0,921
	São João da Urtiga	-0,581	-0,157	-0,522	0,947
	Tupanci do Sul	-0,808	0,553	0,108	0,937
	Campos Borges	-0,611	0,281	0,251	-0,722
	Gramado Xavier	-0,533	-0,549	-0,804	-1,904
	Ibirapuitã	-0,415	-0,243	-0,584	-1,139
	Itapuca	-0,712	0,155	-1,006	-1,266
Lagoão	-0,550	-0,585	-1,357	-2,346	
Mormaço	-0,623	0,205	0,339	-0,711	
São José do Herval	-0,848	0,640	-0,270	-1,473	

* - Não rotacionando as componentes principais.

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

Na segunda etapa, envolvendo a decisão do número de fatores necessários para representar o conjunto de dados, segue-se o critério da raiz latente ou autovalores, pois, segundo Zambrano & Lima (2004, p. 559) “extraem-se apenas os fatores a associados às raízes características maiores que 1” descartando os menores que a unidade. A Figura 10 representa os autovalores graficamente.



Fonte: Figura/output fornecido pelo *software* Statistica 7.0.

Figura 10 – Representação gráfica dos autovalores.

A Tabela 17, a fim de complementar as informações já contidas na figura anterior, apresenta as raízes características, bem como, a percentagem de variância explicada pelos fatores – não rotacionados.

Tabela 17 – Raízes características – *eigenvalues*/autovalores – da matriz de correlação simples (17 x 17) para os determinados COREDEs – 1995/96.

Fatores	Raiz Característica	% Total da Variância	Raiz Característica Acumulada	% da Variância Acumulada
1	9,044	53,200	9,044	53,200
2	2,562	15,068	11,606	68,268
3	1,469	8,639	13,074	76,907
4	1,098	6,458	14,172	83,365
5	0,919	5,409	15,092	88,774
6	0,645	3,797	15,737	92,571
7	0,350	2,056	16,087	94,627
8	0,271	1,594	16,358	96,222
9	0,224	1,318	16,582	97,540
10	0,185	1,088	16,767	98,628
11	0,115	0,679	16,882	99,306
12	0,059	0,345	16,941	99,651
13	0,025	0,146	16,966	99,797
14	0,016	0,096	16,982	99,893
15	0,011	0,062	16,992	99,955
16	0,006	0,032	16,998	99,987
17	0,002	0,013	17,000	100,000
Total	17	100 %		

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

De pronto, os quatro autovalores em destaque são os que melhor representam o conjunto de variáveis e, conjuntamente, explicam aproximadamente 83,37% da variância do conjunto. Dessa forma, opta-se por quatro fatores explicativos para as variáveis originais.

O primeiro fator, isto é, a combinação linear das variáveis originais que pode explicar individualmente a maior parcela da variância, captou aproximadamente 53,20% da variância; o segundo fator representa 15,07% da variância; o terceiro 8,64%; e 6,46% para o quarto, o que totaliza os 83,37% da variância captada pelos quatro autovalores.

Como salientado, o método *Varimax* objetiva melhorar a interpretação dos fatores. Realizando a rotação nos quatro fatores escolhidos têm-se:

Tabela 18 – Percentagem da variância explicada por cada fator após a rotação ortogonal.

Fatores	Antes da rotação (%)	Rotacionado (%)
1	53,200	38,523
2	15,068	21,874
3	8,639	15,044
4	6,458	7,926
Total	83,365	83,365

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

Efetuada a rotação, fica evidente que o fator 1 passou a captar uma proporção menor da variância total das variáveis originais. Entretanto, o poder explicativo do fator 2 e 3 melhoraram significativamente, da mesma forma que o fator 4 teve seu poder de explicação elevado, mas em menor intensidade que nos fatores 2 e 3. Assim, a rotação equilibra o poder de explicação sem alterar o total da variância explicado, que permanece igual ao valor antes da rotação.

Na Tabela 19 são apresentadas as cargas fatoriais ou os coeficientes de correlação entre cada fator e cada uma das 17 variáveis. Também, está a proporção de variância das variáveis originais explicadas por cada um dos cinco fatores, após a rotação pelo método *Varimax*.

Tabela 19 – Cargas fatoriais, comunalidades e percentagem da variância explicada por cada fator após a rotação, na extração de quatro fatores em 17 variáveis.

Indicadores de Modernização	Cargas fatoriais/Coefficiente de Correlação				Comunalidade
	Fator 1	Fator 2	Fator 3	Fator 4	
POP	-0,022	0,568	-0,003	0,573	0,608
PEA	0,508	0,801	-0,090	-0,052	0,952
ASSPEC	0,186	0,755	0,292	0,005	0,665
ASSAGR	0,001	0,799	0,215	0,374	0,753
TRATOR	0,628	0,461	0,540	0,161	0,984
ADUBO	0,179	0,928	-0,039	-0,060	0,951
IRRIGA	0,757	0,150	0,145	0,255	0,860
RECVEG	0,736	0,239	0,519	0,091	0,994
RECANI	0,903	0,247	0,131	0,002	0,942
FINANC	0,520	-0,077	0,672	0,145	0,701
MAQCOL	0,506	0,439	0,654	0,139	0,978
VALVEG	0,741	0,297	0,493	0,084	0,994
VALANI	0,945	0,189	0,140	0,073	0,993
AREXPL	0,963	0,077	0,133	0,074	0,992
RENDA	0,011	0,040	0,807	-0,059	0,471
DISTRG	-0,248	0,004	-0,042	-0,842	0,463
GADO	0,970	0,042	0,060	0,061	0,991
% da variância após a rotação	38,523	21,875	15,043	7,926	

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

Os valores absolutos em destaque foram adotados conforme Hair Jr. (2005, p. 107), ou seja, superiores a 0,50, em virtude do tamanho da amostra – 139 municípios. Dessa forma, as cargas fatoriais fornecem informações para uma interpretação dos fatores ou um significado de cada um, ou seja, é possível batizar os fatores.

O fator 1 associa-se forte e positivamente com o número de tratores (TRATOR – 0,628), valor das receitas advindas de produtos vegetais e animais (RECVEG – 0,736 e RECANI – 0,903), valor dos financiamentos (FINANC – 0,520), número de máquinas para colheita (MAQCOL – 0,506), valor da produção vegetal e animal (VALVEG – 0,741 e VALANI – 0,945), área explorada nos municípios (AREXPL – 0,963), efetivo bovino (GADO – 0,970), PEA (0,508) e, por fim, a variável que mede irrigação (IRRIGA – 0,757). O fator 1 tem uma relação muito forte com as variáveis ligadas ao uso intensivo da terra como GADO, AREXPL, RECANI, VALANI, IRRI e RECVEG. Sem dúvida além de medir a pecuária de corte, o fator mede a modernização da agricultura, indicando uma forte relação com os insumos modernos, denotando um caráter capital-intensivo. Dessa forma, batiza-se o fator 1 como o fator que mede a “agricultura moderna”, fortemente ligada aos insumos tecnológicos e pecuária.

Observa-se que o fator 2 encontra-se fortemente correlacionado com a população residente no campo (POP – 0,568), novamente, a população economicamente ativa (PEA – 0,801), contudo mais fortemente ligada a este fator, assistência técnica especializada na produção animal (ASSPEC – 0,755), assistência técnica especializada na produção vegetal (ASSAGR – 0,799) e número de estabelecimentos que utilizam adubos químicos (ADUBO – 0,928). Em suma, o fator 2 pode ser interpretado como o fator que mede a agricultura e o apoio a inovação tecnológica e químico-biológica, além de denotar um caráter trabalho-intensivo. O fator 2 fica batizado como o fator “apoio à inovação”.

Dessa forma, os fatores anteriores apresentam uma semelhança, pois ambos estão medindo a modernização na agricultura. Sendo que, o segundo está captando o fenômeno mais na dimensão do apoio à inovação da agricultura e, o primeiro fator, capta mais em termos de insumos modernos.

O fator 3 apresenta correlação positiva e forte com o número de tratores (TRATOR – 0,540), valor das receitas advindas de produtos vegetais (RECVEG – 0,519), valor dos financiamentos (FINANC – 0,672), número de máquinas para colheita (MAQCOL – 0,654) e com o índice de renda dos municípios (RENDA – 0,807). A correlação significativa das variáveis TRATOR, RECVEG, MAQCOL e FINANC também se repetem no fator 1. Contudo, como os financiamentos são gerados para aquisição de maquinário pesado –

colheitadeiras, por exemplo – , e, estas, são utilizadas para cobrir grandes extensões de terra – monocultura exportadora da soja, por exemplo – , neste fator, leva-se em consideração apenas a variável FINANC dentre as quatro supra-expostas. Assim, RENDA e FINANC caracterizam o fator 3, e este é batizado como o fator que mede o “nível de renda e valores financiados dos municípios” o que, de certa forma, está ligado a uma agricultura moderna.

Por fim, o fator 4: único, com associações negativas, tem uma correlação forte e negativamente com a variável distância do porto de Rio Grande (DISTRG – -0,842) e com a população rural (POP – 0,573). Sendo evidente, que este mede a relação espacial dos municípios, em relação à distância do porto de Rio Grande, por onde escoam a maior parte da produção agropecuária exportada gaúcha. Dessa forma esse fator assume a função “espacial” que é um componente fundamental na competitividade.

Como o objetivo da técnica de análise multivariada é uma redução de dados para facilitar a interpretação/conclusões acerca de fenômenos, em especial, a Análise Fatorial e de Agrupamentos, busca-se simplificar a matriz de dados $X_{139 \times 17}$. De fato, a análise de *cluster* se deteve em reduzir o número de observações – as linhas da matriz – enquanto a Análise Fatorial reduziu o número de variáveis – colunas. Foram criadas 4 novas variáveis – fatores – que juntas explicam aproximadamente 83,37% da variância do conjunto.

A Tabela 20 apresenta a relação dos fatores com todos os municípios do estudo, mas, como previsto fica complexo de analisar caso a caso, em vista da quantidade de observações.

Tabela 20 – Escores fatoriais para os grupos de I a IV após rotacionados os fatores.

Grupos	Municípios	Fator 1	Fator 2	Fator 3	Fator 4
Grupo I	Alegrete	5,280	0,777	-0,488	-0,974
	Itaqui	3,119	-0,901	3,230	-1,173
	Rosário do Sul	2,799	-0,321	-0,866	0,726
	Santana do Livramento	4,296	-0,457	-2,422	0,233
	São Borja	1,672	-0,237	2,788	-0,036
	São Gabriel	4,264	-0,395	-0,335	1,345
	Uruguaiana	4,525	-0,200	2,753	-0,671
	Cachoeira do Sul	2,957	2,278	-0,355	1,993
Grupo II	Giruá	-0,028	3,151	1,071	-0,514
	Ajuricaba	-0,271	1,380	0,809	-0,012
	Augusto Pestana	-0,405	1,265	0,527	-0,045
	Crissiumal	-0,289	3,164	-0,700	-1,622
	Ijuí	-0,139	2,859	1,266	-0,047
	Panambi	-0,522	1,232	1,171	0,066
	Três Passos	-0,197	3,223	-0,823	-1,768
	Sananduva	-0,074	1,703	0,007	-1,216
	Tapejara	-0,196	1,624	0,326	-0,963
	Cruz Alta	0,145	3,194	2,720	1,800
	Ibirubá	-0,737	2,338	1,565	1,434
	Barros Cassal	-0,757	2,269	-1,841	2,903
	Espumoso	-0,725	2,203	0,397	1,399
	Fontoura Xavier	-0,615	1,886	-1,890	2,899
	Soledade	-0,054	0,921	-0,846	1,563

Tabela 20 – Escores fatoriais para os grupos de I a IV após rotacionados os fatores.

(Continua)

Grupos	Municípios	Fator 1	Fator 2	Fator 3	Fator 4
Grupo III	Quarai	1,501	-1,234	-0,592	0,264
	Cerro Branco	-0,372	0,002	-0,735	1,954
	Paraíso do Sul	-0,183	-0,018	-0,634	2,345
	Restinga Seca	0,516	-0,032	0,138	2,598
	São Sepé	1,252	0,211	0,275	2,297
	Vila Nova do Sul	-0,123	-1,665	-0,426	2,182
	Bossoroca	0,467	-0,728	0,464	0,143
	Caibaté	-0,221	0,527	-0,215	-0,516
	Cerro Largo	-0,520	0,534	0,245	-0,591
	Entre-Ijuís	-0,172	0,072	0,617	-0,124
	Eugênio de Castro	-0,329	-0,561	0,433	-0,009
	Garruchos	-0,135	-1,030	0,727	-0,994
	Guarani das Missões	-0,305	0,946	-0,341	-0,404
	Porto Xavier	-0,206	0,191	-0,558	-1,093
	Roque Gonzáles	-0,094	0,356	-0,514	-0,722
	Salvador das Missões	-0,788	-0,257	0,958	-0,477
	Santo Ângelo	-0,177	0,953	0,201	-0,145
	Santo Antônio das Missões	1,009	-0,595	-0,445	-0,292
	São Luiz Gonzaga	0,748	0,986	0,275	-0,670
	São Miguel das Missões	0,288	-0,699	1,836	-0,028
	São Paulo das Missões	-0,092	0,493	-0,873	-0,842
	Bom Progresso	-0,558	-0,957	0,469	-0,416
	Campo Novo	-0,403	-0,925	0,740	-0,151
	Catuípe	0,032	0,009	0,478	0,007
	Chiapeta	-0,364	-0,751	0,986	-0,095
	Condor	-0,393	-0,365	1,297	0,240
	Coronel Barros	-0,587	-0,797	0,883	0,161
	Coronel Bicaco	-0,296	-0,142	0,676	-0,274
	Inhacorá	-0,539	-1,174	0,701	-0,257
	Jóia	-0,145	0,268	1,122	0,186
	Pejuçara	-0,450	-0,784	1,698	0,510
	Santo Augusto	-0,369	-0,477	1,234	-0,067
	São Martinho	-0,344	0,260	-0,033	-0,555
	Tenente Portela	-0,240	0,760	-0,398	-0,751
	Tiradentes do Sul	-0,102	0,732	-0,735	-1,193
	Alpestre	-0,053	1,301	-1,605	-0,988
	Boa Vista das Missões	-0,648	-1,203	1,359	-0,197
	Dois Irmãos das Missões	-0,518	-1,136	1,322	-0,491
	Engenho Velho	-0,733	-0,565	0,819	-0,454
	Ervál Seco	-0,054	0,290	-0,962	-0,434
	Frederico Westphalen	-0,341	0,654	-0,062	-0,448
	Gramado dos Loureiros	-0,515	-1,118	0,705	-0,778
	Seberi	-0,230	0,621	-0,571	-0,263
	Água Santa	-0,035	0,153	-0,182	-0,682
	Ibiaçá	-0,308	0,348	0,341	-0,670
	Ibiraiaras	-0,339	0,534	0,274	-0,734
	Lagoa Vermelha	0,756	0,905	-0,415	-1,092
São José do Ouro	-0,425	0,269	0,440	-0,942	
Colorado	-0,768	-0,082	1,249	0,957	
Fortaleza dos Valos	-0,375	-0,482	1,648	0,847	
Lagoa dos Três Cantos	-0,631	-0,732	0,810	0,494	
Não-Me-Toque	-0,617	0,374	1,389	0,724	
Quinze de Novembro	-0,497	0,055	0,339	0,747	
Saldanha Marinho	-0,615	-0,697	1,008	0,540	
Salto do Jacuí	-0,186	-0,608	0,568	1,267	
Santa Bárbara do Sul	-0,194	-0,421	2,322	1,164	
Selbach	-0,703	0,158	0,771	0,938	
Tapera	-0,627	-0,558	1,537	0,893	
Alto Alegre	-0,429	-1,177	-0,518	3,831	
Nicolau Vergueiro	-0,506	-0,972	0,763	0,267	
Victor Graeff	-0,854	0,423	1,323	0,508	

Tabela 20 – Escores fatoriais para os grupos de I a IV após rotacionados os fatores. (Continua)

Grupos	Municípios	Fator 1	Fator 2	Fator 3	Fator 4
Grupo IV	Itacurubi	0,602	-1,156	-0,747	0,189
	Manoel Viana	0,321	-1,127	0,187	0,011
	Dezesseis de Novembro	-0,041	-0,725	-1,083	-0,467
	Pirapó	-0,143	-0,555	-0,650	-0,707
	São Nicolau	0,055	-0,706	-0,481	-0,514
	São Pedro do Butiá	-0,388	-0,286	0,139	-0,639
	Vitória das Missões	-0,252	-0,480	-0,236	-0,181
	Barra do Guarita	-0,389	-1,028	-0,039	-0,807
	Braga	-0,336	-0,769	-0,129	-0,242
	Derrubadas	-0,398	0,061	-0,023	-0,905
	Humaitá	-0,464	-0,066	0,155	-0,569
	Miraguaí	-0,076	-0,251	-1,168	-0,555
	Redentora	-0,122	-0,066	-0,883	-0,191
	São Valério	-0,193	-0,913	-0,701	-0,038
	Sede Nova	-0,449	-0,458	0,141	-0,480
	Vista Gaúcha	-0,280	-0,115	-0,555	-0,658
	Ametista do Sul	-0,118	-1,043	-1,052	-0,059
	Caiçara	-0,349	0,337	-0,788	-0,069
	Cerro Grande	-0,123	-0,696	-1,081	-0,119
	Iraí	-0,308	-0,039	-0,492	-0,601
	Jaboticaba	-0,115	-0,612	-0,892	0,077
	Lajeado do Bugre	-0,002	-1,154	-1,296	0,195
	Liberato Salzano	-0,185	0,231	-1,006	-0,595
	Nonoai	-0,256	-0,420	0,231	-0,802
	Novo Tiradentes	-0,193	-0,801	-0,885	-0,020
	Palmitinho	-0,256	0,351	-0,931	-0,366
	Pinhal	-0,309	-0,923	-0,435	-0,055
	Pinheirinho do Vale	-0,263	-0,129	-0,664	-0,374
	Planalto	-0,074	0,031	-1,032	-0,813
	Rio dos Índios	-0,369	-0,171	-0,016	-0,906
	Rodeio Bonito	-0,327	0,078	-0,629	-0,366
	Sagrada Família	-0,249	-0,990	-0,480	-0,275
	Taquaruçu do Sul	-0,426	-0,754	0,069	-0,250
	Três Palmeiras	-0,285	-0,592	-0,066	-0,561
	Trindade do Sul	-0,146	-0,161	-0,542	-0,741
	Vicente Dutra	-0,055	0,042	-1,384	-0,318
	Vista Alegre	-0,275	-0,600	-0,588	-0,204
	André da Rocha	-0,006	-1,451	-0,482	0,477
	Barracão	-0,111	-0,187	-0,248	-1,049
	Cacique Doble	-0,278	0,496	-0,963	-0,761
	Caseiros	-0,102	-0,853	-0,527	-0,424
	Esmeralda	0,536	-0,566	-1,294	-0,435
	Machadinho	0,140	-0,050	-1,158	-1,184
Maximiliano de Almeida	-0,133	-0,040	-0,606	-1,224	
Paim Filho	-0,187	-0,132	-0,356	-1,121	
Santo Expedito do Sul	-0,320	-0,191	-0,463	-1,055	
São João da Urtiga	-0,308	0,138	-0,596	-1,030	
Tupanci do Sul	-0,398	-0,659	-0,049	-1,119	
Campos Borges	-0,477	-0,686	-0,065	0,577	
Gramado Xavier	-0,442	-0,032	-1,211	1,787	
Ibirapuitã	-0,262	-0,122	-0,848	1,033	
Itapuca	-0,160	-0,569	-1,327	1,016	
Lagoão	-0,279	-0,026	-1,779	2,179	
Mormaço	-0,555	-0,635	0,005	0,583	
São José do Herval	-0,329	-1,184	-0,691	1,176	

Fonte: Resultados da pesquisa obtidos através do *Statistica*.

Dessa forma, como os 139 municípios foram agrupados em 4 grupos, o mais homogêneos possível, é correto inferir que os grupos formados representam a realidade dos municípios neles contidos. Sendo assim, a antiga matriz com 139 linhas e 17 colunas passa a ser agora representada por uma nova, reduzida, com quatro linhas e quatro colunas. Esta pode ser vislumbrada abaixo:

$$X_{4 \times 4} = \begin{bmatrix} 3,614 & 0,068 & 0,538 & 0,180 \\ -0,324 & 2,161 & 0,250 & 0,392 \\ -0,215 & -0,157 & 0,404 & 0,118 \\ -0,200 & -0,414 & -0,586 & -0,268 \end{bmatrix}$$

Dessa forma tem-se a nova matriz gerada após a aplicação da técnica multivariada, e, conseqüente e pretendida redução do número inicial de dados sem que se perca muito das informações contidas na mesma. Isto só foi possível pela complementação de uma técnica a outra, ou seja, agrupando-se primeiro para logo após explicar as razões deste agrupamento pelos fatores gerados na Análise Fatorial. Isso pode ser melhor visualizado na Tabela 21.

Tabela 21 – Classificação em ordem decrescente do nível tecnológico, participação percentual no valor da produção agropecuária, animal e vegetal para os quatro grupos homogêneos dos COREDEs agrícolas do Estado do Rio Grande do Sul.

Grupo	Fator 1	Fator 2	Fator 3	Fator 4	Número de Municípios	% VAB* agropecuária	%VBP animal	%VBP vegetal
I	3,614	0,068	0,538	0,180	8	5,76%	9,36	11,85
II	-0,324	2,161	0,250	0,392	15	10,79%	6,65	7,20
III	-0,215	-0,157	0,404	0,118	61	43,88%	13,77	15,28
IV	-0,200	-0,414	-0,586	-0,268	55	39,57%	6,23	5,42
					n	139	100%	

Fonte: Resultados da pesquisa obtidos através do *Statística*.

* Valor adicionado bruto agropecuário dos municípios em 1995, dado fornecido pela Fundação de Economia e Estatística – FEE.

Assim, depreende-se que os municípios que formam o grupo I (escore de 3,614) apresentam uma moderna e dinâmica agricultura voltada para o comercial – indústria e exportação – e exploração da pecuária conjuntamente pelo fato de terem o maior escore no fator 1. Esse escore coaduna com a realidade dos municípios do grupo, pois, são zonas tradicionais de produção de produtos do agronegócio gaúcho. Em um patamar inferior ao

nível de modernização, estão os grupos II, III e IV, respectivamente. Sendo que todos os grupos demonstraram uma ausência de escore positivo no fator agricultura moderna e pecuária de corte, denotando um caráter de agricultura mais tradicional, baseada em minifúndios e pequenas propriedades, ou seja, não voltada para latifúndios agro-exportadores intensivos em capital.

Os resultados pela análise multivariada são verossímeis, pois, o grupo I formado de apenas oito municípios é responsável por 9,36% do VAB do Rio Grande do Sul (RS), ainda, representa 14,55% e 11,85% do valor bruto da produção animal de grande porte e vegetal do RS. Em trabalho semelhante, Freitas, Paz & Nicola (2007) chegaram a um resultado proporcionalmente semelhante, ou seja, um pequeno grupo representa uma parcela importante do valor de produção. O grupo II responde por 6,65% do VAB do Estado e 6,01% e 7,20% da produção animal e vegetal, respectivamente. O grupo III tem uma representatividade significativa, pois VAB responde por 13,77% do RS, mas, lembrando que a representatividade bem superior aos demais, deve-se ao fato de este grupo conter aproximadamente 44% da amostra selecionada para o estudo. O grupo IV, apesar de tamanho elevado, 39,57% da amostra, responde por 6,23% do VAB do Estado, o que, corrobora a sua relação com os fatores 1 e 2. O grupo IV tem uma relação inversa à modernização e, também, com o apoio à inovação.

A relação do grupo I com o fator 2 é mais modesta, que mede a agricultura subjacente menos intensiva em capital e que responde por grande parte das receitas vegetais, ratificando o elevado potencial de modernização já atingido por estes. Dessa forma, como salientado, é o grupo com o maior valor bruto agrícola e pecuário.

Já o grupo II, pelo seu escore no fator 1 (-0,324) e fator 2 (2,161) denota uma agricultura também dinâmica tendo um apoio à modernização tecnológica e químico-biológica, vide sua relação com as variáveis que mensuram os estabelecimentos com assistência técnica e utilização de adubos. Pese o fato de concentrar, ainda, um grande número de mão-de-obra empregada.

Já os grupos III e IV, não possuem uma estrita ligação com caráter de agricultura moderna e pecuária extensiva. Comparado aos grupos I e II, os grupos III e IV, demonstram ser os grupos que contêm a produção menos dinâmica. Depreende-se que os elementos do grupo III, apesar do baixo escore para com o fator 2 (-0,157) e com o fator 1 (-0,215), se comparado ao fator 4, demonstram um padrão de transição para uma agricultura mais dinâmica apoiada na modernização, pois, estão no intermediário entre os grupos II e IV. Ainda, este grupo é o maior da análise, pese o fato de ainda não ter uma agricultura moderna

ou apoiada. Em termos de potencial, a situação do grupo III é superior a do grupo IV, sendo passível de políticas públicas específicas.

Quanto aos escores para os grupos referentes ao fator 3 – nível de renda e financiamentos – pode inferir-se que: a) o grupo I coadunando a elevada representatividade econômica apresenta o mais elevado nível de renda e valores em financiamentos (0,538), pois este responde por aproximadamente 37% dos financiamentos da amostra e 18% de todo Estado; b) o grupo II para esse fator está mais abaixo em relação ao grupo III (0,250); c) o grupo III apresenta valores mais elevados que o grupo II (0,404), reforçando a idéia de transição, pois, se verifica uma renda mais elevada e valores de financiamentos superiores que o já estabelecido grupo II; d) em virtude de todo histórico revelado pelos fatores 1 (-0,200), 2 (-0,414) e 3 (-0,586) para o grupo IV, este se encontra em pior situação socioeconômica em relação aos demais e sem demonstração de uma reação própria para isto. Por se tratar de COREDE agrícola de baixa correlação com fatores 1 e 2, assim, não tem elevados valores em financiamentos de implementos, reforçando o caráter mais tradicional da produção.

E, por fim, o grupo II é o grupo de melhor relação com o fator espacial, possuindo o escore mais elevado (0,392). O grupo III encontra-se, como o grupo I, também, com boa ligação com o porto, pois, o grupo I obteve escore de (0,180), não tão alto quanto o grupo II, uma vez que, ele é composto de municípios que na sua maioria, estão localizados na fronteira do Estado com Argentina e Uruguai. O grupo III obteve escore de (0,118) seguido do grupo IV, que novamente, deteve o pior escore dos quatro grupos – (-0,268). Estando formado pelos municípios com maior distância do porto de Rio Grande. Assim, considerando que a distância do porto de Rio Grande é um fator de competitividade e, levando em conta que 25% da produção se perde, antes mesmo, de chegar ao embarque no porto (MELO, 1990, p. 61), o grupo 2 apresenta uma vantagem competitiva em relação aos demais.

Finalizando a análise, considerando as informações quantitativas dos municípios e suas realidades/peculiaridades, este trabalho aponta para uma forte concentração da agricultura moderna e da agropecuária no grupo I. Uma situação mais estável para o grupo II e, uma transição para o grupo III que demonstrou sinais de transição da agricultura tradicional para uma mais apoiada na inovação. No grupo III, com escore (-0,157) para o fator 2, aproximadamente 48% dos municípios tem escores positivos com este fator, reforçando a idéia de transição do grupo, em especial, dessa parcela de municípios. Observa-se que o grupo IV tem o pior grau de modernização da agricultura do Estado e sem exploração intensiva da área para pecuária. Em virtude de seu elevado tamanho, carece de uma atenção especial para

que eventuais potencialidades possam ser exploradas, mesmo que este apresente um padrão de agricultura tradicional.

6 CONCLUSÕES

Tendo por objetivo, agrupar municípios o mais homogêneos possível, tanto a Análise de Agrupamentos, como a Análise Fatorial, mostraram seu importante papel para a pesquisa econômica. O objetivo principal de mensurar o grau de modernização dos municípios, em vista de seu elevado número, foi possível pela utilização das técnicas de análise multivariada de dados socioeconômicos, revelando, assim, uma excelente ferramenta para pesquisas no ramo da Economia e Desenvolvimento Rural.

Nos resultados obtidos, a realidade da modernização brasileira – excludente e conservadora – se repetiu no Estado, pois numa minoria de municípios, oito, dispõem de elevado grau de modernização na agricultura, bem como, são responsáveis por aproximadamente 14,55% do Valor Bruto da Produção Animal. Essa realidade demonstra a elevada concentração da agricultura, dita moderna, no grupo I.

Por outro lado, revelou que 131 municípios da amostra não têm uma agricultura considerada moderna. O grupo II de tamanho médio, com 15 municípios, não tem o padrão tecnológico encontrado no grupo I, mas tem uma agricultura ainda dinâmica e com apoio à modernização tecnológica e químico-física. Ainda, revela que o grupo III, apesar de seu elevado tamanho, está em um estágio que pode ser entendido de transição entre a agricultura mais tradicional, para uma apoiada na modernização e em insumos mais modernos, dessa forma, tomando um caráter mais dinâmico na agropecuária. O resultado mais preocupante é sobre o grupo IV, que contém 55 municípios, apresentando a pior relação com os fatores de agricultura moderna – fato 1– e fator de apoio à inovação tecnológica – fator 2.

Quanto ao nível de renda e financiamentos, novamente se destaca o grupo I – que, em oito municípios, concentram 37% do valor dos financiamentos – seguido dos grupos III e II, respectivamente. O grupo III tem resultado maiores que o grupo II, reiterando a sua condição de transição a uma agricultura mais dinâmica e apoiada em tecnologia na agricultura. Novamente, o grupo IV é o de pior resultado para financiamentos e renda.

O fator espacial, que buscou interagir entre a distância do porto e da cidade produtora, se mostrou eficiente. De modo que, se a competitividade está relacionada a menores custos de transporte também, os grupos II, I e III se mostraram, conforme a ordem, competitivos, o que não se verificou no grupo IV mais uma vez.

Com base no estudo, infere-se que os municípios do grupo I são os que têm melhor nível de modernização da agricultura e pecuária, demonstrando, em virtude de suas longas

histórias de fronteiras agrícolas gaúcha, consolidados no cenário do Rio Grande do Sul. O grupo II encontra-se num padrão dinâmico de agropecuária, seguido do grupo III que, no período estudado, demonstra uma transição entre agricultura tradicional e dinâmica.

Por fim, observa-se que o grupo IV tem o pior grau de modernização da agricultura do Estado e sem exploração intensiva da área tanto para agricultura como para pecuária. Em virtude de seu elevado tamanho, carece de uma atenção especial para que eventuais potencialidades possam ser exploradas, mesmo que, este apresente um padrão de agricultura tradicional.

Para trabalhos futuros, após a publicação do Censo Agropecuário 2007/08, sugere-se um novo retrato da modernização agrícola no Rio Grande do Sul. Para com isso, comparar se persiste a concentração da modernização nas áreas do grupo I e, se, os municípios do grupo III conseguiram ingressar em um grupo mais dinâmico.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, Francisco; PAULILLO, Luiz F. **Crise Agrícola e Políticas Públicas: Novos Elementos para Discussão**. Agricultura em São Paulo, SP, 45(2): 17-38. 1998.

BACHA, C. J. C. **Economia e Política Agrícola no Brasil**. São Paulo: Atlas, 2004.

BANCO CENTRAL DO BRASIL – BC. Disponível em: <<http://www.bcb.gov.br>>. Acesso em: 25 mar. 2004.

BARROS, G. S. de C.; ARAÚJO, P. F. C. de. **Crédito Rural no Brasil: estrutura de oferta e demanda e perspectivas para os anos 90**. Piracicaba: ESALQ-USP, 1995. (Série Pesquisa n. 49.)

BNDES: FINAME Agrícola. Disponível em: <<http://www.bndes.gov.br/atuar/finameag.htm>>. Acesso em: nov. 2003.

BOUROCHE, J. N.; SAPORTA, G. **Análise de dados**. Rio de Janeiro: Zahar, 1982.

BRUM, A. J. **O desenvolvimento econômico brasileiro**. 19 ed. Ijuí (RS): Vozes/Unijuí, 1998.

_____. **O Desenvolvimento Econômico Brasileiro**. 4 ed. Petrópolis: Vozes, 1984.

BRUM, A. L. Economia da soja: História e Futuro. **AgroBr: Artigos**. Disponível em: <<http://www.agrobr.com/>>. Acesso em: 10 mar. 2004.

CEPEA. Centro de Estudos Avançados em Economia Aplicada. Disponível em: <<http://www.cepea.esalq.usp.br>>. Acesso em: 10 jun. 2007.

COELHO, C. N. A Lei Agrícola Americana de 2002 e o Comércio Mundial. **Revista de Política Agrícola**, abr./jun. 2002.

CONAB. Companhia Nacional de Abastecimento. Disponível em: <<http://www.conab.gov>>. Acesso em: 25 jul. 2007.

DIAS, G. L. da S. **Mobilização de poupança e o financiamento do setor agrícola**. Piracicaba: ESALQ-USP, 1995. (Série Pesquisa n. 51).

EMATER. Associação Riograndense de Empreendimentos de Assistência Técnica e Extensão Rural. Disponível em: <<http://www.emater.tche.br>>. Acesso em: 25 mar. 2007.

EMBRAPA. Empresa Brasileira de Pesquisa Agropecuária. Disponível em: <<http://www.embrapa.gov.br>>. Acesso em: 30 jun. 2007.

FARSUL. Acesso em: < http://www.farsul.org.br/pg_informes.php?id_noticia=6883 >. Acesso em: 34 jul. 2004.

FAGUNDES, M. H. **Comentários sobre crédito rural e sua evolução recente**. Brasília: CFP, 1986. (Coleção de Estudos Especiais, v.21).

FREITAS, C.; PAZ, M. V.; NICOLA, D. S. Avaliando a intensidade da modernização da agropecuária gaúcha: uma aplicação de análise fatorial e cluster. **Revista Análise Econômica**. Porto Alegre, ano 25, n. 47, p. 121-149, 2007.

FREITAS, N. **Moratória: Moratória Mexicana e Brasileira**. Disponível em: <<http://www.newton.freitas.nom.br/artigo092.php>>. Acesso em: 25 mai. 2004.

FUNDAÇÃO DE ECONOMIA E ESTATÍSTICA – FEE. Idese 1991-2000. Disponível em: <<http://www.fee.tche.br>>. Acesso em: 10 jun. 2007.

GAMA, H. O renascimento do transporte hidroviário no RS. **Voto**, Ano 3, n. 35, p. 26 – 29, 2007.

GASQUES, J. G.; SPOLADOR, H. F. S. Taxa de Juros e Políticas de apoio Interno à Agricultura. Brasília: IPEA, abr. 2003. (Texto para discussão, 952).

GASQUES, J. G.; VILLA VERDE, C. M. Recursos para a agricultura e orientação dos gastos públicos. In: **IPEA. Perspectivas da Economia Brasileira – 1992**. Brasília: IPEA, 1991.

_____. Crescimento da Agricultura Brasileira e Política Agrícola nos anos oitenta. **Agricultura em São Paulo**, São Paulo, v.37, n.1, 1990.

_____. **Gastos Públicos na Agricultura**. Brasília: IPEA, mar. 2001. (Texto para discussão, 782).

_____. **Prioridade e Orientação dos Gastos Públicos em Agricultura no Brasil**. Brasília: IPEA, jan. 1995. (Texto para discussão, 365).

_____. Gastos Públicos na Agricultura: Evolução e Mudanças. **Bahia Análise & Dados**, Salvador, v.12, n.4, p.133-154, mar. 2003.

_____. **Crescimento e Produtividade da Agricultura Brasileira**. Brasília: IPEA, jul. 1997. (Texto para discussão, 502).

GRAZIANO DA SILVA, J. **A Modernização Dolorosa**. Zahar, Rio de Janeiro, 1982.

_____. **Tecnologia e Agricultura Familiar**. Campinas, SP: UNICAMP, 1999.

_____. **A nova Dinâmica da Agricultura Brasileira**. Campinas, SP: UNICAMP, 1996.

_____. A Gestão das Políticas na Agricultura Brasileira Moderna. In: CONGRESSO DE ECONOMIA E SOCIOLOGIA RURAL, Brasília, vol. 27, n. 3, 1989. **Anais...**Brasília: SOBER, jul./ago./set. 1989 p. 309-326.

GRAZIANO, X. O Conceito de Latifúndio. **O Estado de São Paulo**, São Paulo, 15 abr. 2003.

GREMAUD, Amaury Patrick; VASCONCELLOS, Marco Antonio S.; TONETO, Rudinei Toneto. **Economia Brasileira Contemporânea**. 4. ed. São Paulo: Atlas S. A., 2002.

GUIA 4 RODAS. Disponível em: <<http://www.guia4rodas.abril.com.br/mapasbrasi>>. Acesso em: 10 abr. 2007.

HAIR, Jr., J. F. **Análise multivariada de dados**. 5ª ed. Porto Alegre: Bookman, 2005. 593 p.

HÄRDLE, W.; SIMAR, L. **Applied Multivariate Statistical Analysis**. Berlin and Louvain-la-Neuve, 2003. E-book disponível em: <<http://www.xplore-stat.de>>.

HOFFMANN, R. **Componentes principais e análise Fatorial**. Viçosa: Série Didática nº 90, 1999. 4ª ed. 40 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 1 abr. 2007.

JOHNSTON, B.F.; MELLOR, J.W. The role of agriculture in economic development. **American Economic Review**, v.51, n.4, p.566-593, Set. 1961.

JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 3ª ed. New Jersey: Prentice Hall, 1992.

LOPES, I. V.; ROCHA, D. P. Agricultura familiar: muitos produzem pouco. In: **Conjuntura Econômica**. Vol. 59, n. 02, p. 30-35. Fev. 2005.

MALHOTRA, N. K. **Pesquisa de Marketing: uma orientação aplicada**. Porto Alegre: Bookman, 2001.

MANLY, B. F. J. **Multivariate statistical methods: a primer**. London: Chapman and Hall, 1986.

MARDIA, B. F.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis**. London: Academic, 1979.

MASSUQUETTI, A. **A Mudança no padrão de financiamento da agricultura brasileira**. Porto Alegre: UFRGS, 1998. 222f. Dissertação (Mestrado em Economia Rural) Universidade Federal do Rio Grande do Sul, 1998.

MELO, F. B. H. O crescimento agrícola brasileiro dos anos 80 e as perspectivas para os anos 90. **Revista de Economia Política**, v.10, n. 3, set. 1990.

MINISTÉRIO DA FAZENDA – BRASIL. Secretaria do Tesouro Nacional. Relatório de Atividades, Vários Anos.

_____. Secretaria do Tesouro Nacional. Informações do Balanço Geral da União. Disponível em: <<http://www.stn.fazenda.gov.br>>. Acesso em: jun. 2004.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005. 279 p.

MONTEIRO, M. J. C. Renda Agrícola: terra sem promessa. **Agroanalysis**, Rio de Janeiro, Fundação Getúlio Vargas, v. 19, n. 4, p. 17-18, abril 1999.

MORRISON, D. F. **Multivariate statistical methods**. 2 Ed., New York: Mc Graw Hill, 1976.

MÜLLER, C. A. **A História Econômica o Rio Grande do Sul**. Porto Alegre: Ed. Grande Sul, 1999.

PRIEB, R. I. P. **Pluriatividade na produção familiar fumageira**. Santa Cruz do Sul: EDUNISC, 2005. 195 p.

RANGEL, I. **Questão agrária, industrialização e crise urbana no Brasil**. Porto Alegre: Editora da UFRGS, 2004.

REGAZZI, A. J. **INF 766 – Análise multivariada**. Viçosa: Universidade Federal de Viçosa. Centro de Ciências Exatas e Tecnológicas. Departamento de Informática, 2001. 166 p. Mimeo.

REIS, E. **Estatística Multivariada**. Lisboa: Edições Silabo Ltda, 1997. 343 p.

RIO GRANDE DO SUL. Rumos 2015: estudo sobre desenvolvimento regional e Logística de Transportes no RS. Maio de 2005. 97 p.

ROMANO, P. A. Investimentos para a Produtividade Agrícola. In: SEMINÁRIO NACIONAL DE PRODUTIVIDADE, Rio de Janeiro, 1987. **Anais...**Rio de Janeiro: CEBRAE, out. 1987, p. 101-110.

RUMMEL, R. J. Understanding factor analysis. **Journal of Conflict Resolution**, Vol. XI, n. 4, dec. 1967. p. 444-480.

POERSCHKE, R. **Implicações das Políticas Governamentais Pós 80 Sobre o Desempenho do Complexo Soja Gaúcho**, RS. 2004. 102 f. Monografia de Graduação (Ciências Econômicas) – Universidade Federal de Santa Maria, Santa Maria, 2004.

PORTO NETO, A. A política e a globalização. **Revista de Política Agrícola**, Brasília, ano 5, n. 4, p. 5-11, out./nov./dez. 1996. (Publicação Trimestral.)

SANDRONI, P. **Novíssimo Dicionário de Economia**. 9. ed. São Paulo: Best Seller, 2002.

SANTOS, R. F. **O Crédito Rural na Modernização da Agricultura Brasileira**. In: CONGRESSO DE ECONOMIA E SOCIOLOGIA RURAL, Brasília, vol. 26, n. 4, 1988. **Anais**.Brasília: SOBER, out./dez. 1998 p. 393-404.

SCHNEIDER, S.; FIALHO, A. A. V. Pobreza rural, desequilíbrios regionais e desenvolvimento agrário no Rio Grande do Sul. **Teoria e Evidência Econômica**. Passo Fundo, v.8, n.15, p. 117-150, nov. 2000.

SOUZA, A. M.; VICINI, L. **Análise Multivariada: da teoria à prática**. Santa Maria: UFSM, CCNE, 2006. 215 p.

SOUZA, N.de J.de; LUCENA, R. B. O papel da agricultura no desenvolvimento econômico brasileiro, 1980/1998. **Análise Econômica**, Porto Alegre, RS, v. 19, n. 35, p. 55-72, mar. 2001.

_____. Políticas agrícolas e desempenho da agricultura brasileira: 1950-00. **Indicadores Econômicos FEE**, Porto Alegre, v. 29, n. 02, p. 180-200, ago. 2001.

TEIXEIRA, G. Um Perfil da Agricultura e da Reforma Agrária no Brasil entre 1995 e 1999. **Partido dos Trabalhadores**, Brasília, mai. 2000. Disponível em: < <http://www.pt.org.br.htm> >. Acesso em: 10 jun. 2004.

VALENTIN, J. . Ecologia numérica - uma introdução à análise multivariada de dados ecológicos. Rio de Janeiro: Interciência, 2000. 117 p.

VICINI, L. **Análise multivariada: da teoria à prática**. Monografia (Especialização) – Universidade Federal de Santa Maria, Centro de Ciências Naturais e Exatas, Programa de Pós-Gduação em Estatística e Modelagem Quantitativa, RS, 2005. 140 p.

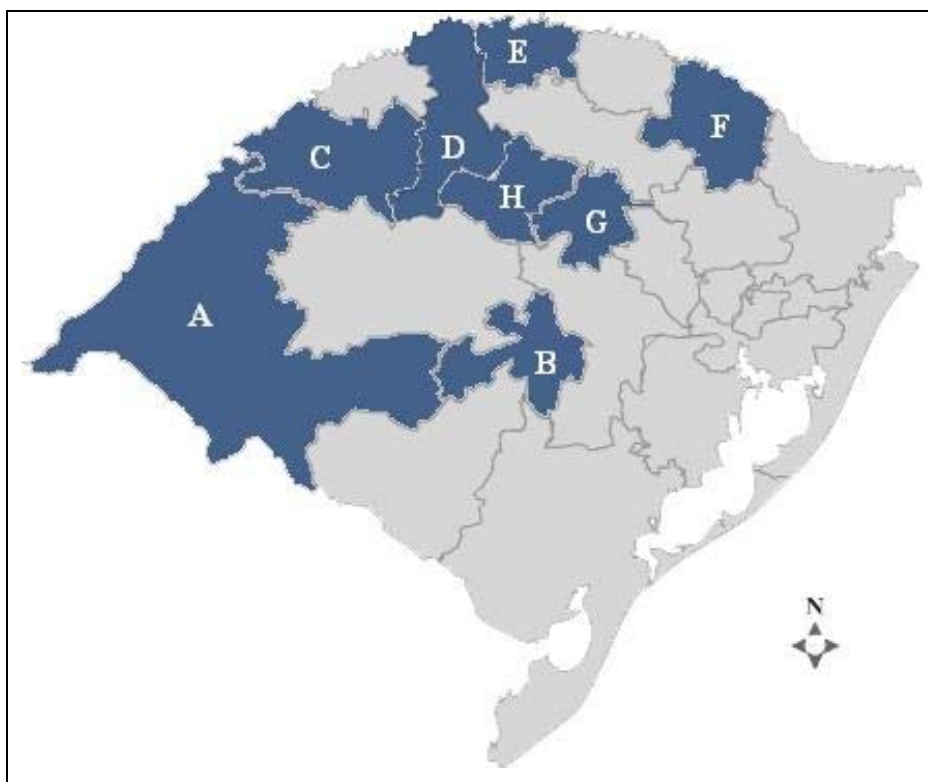
ZAMBONADI, R. **Fundamentos Técnicos para o Diagnóstico da Agricultura Brasileira**. Brasília: CNA, 1996.

ZAMBRANO, C.; LIMA, J. E. Análise Estatística Multivariada de Dados Socioeconômicos. In: SANTOS, M. L.; VIEIRA, W. C. (Ed.). **Métodos quantitativos em economia**. Viçosa, UFV, 2004. p.556-577.

ANEXOS

ANEXO A – COREDEs utilizados para o estudo

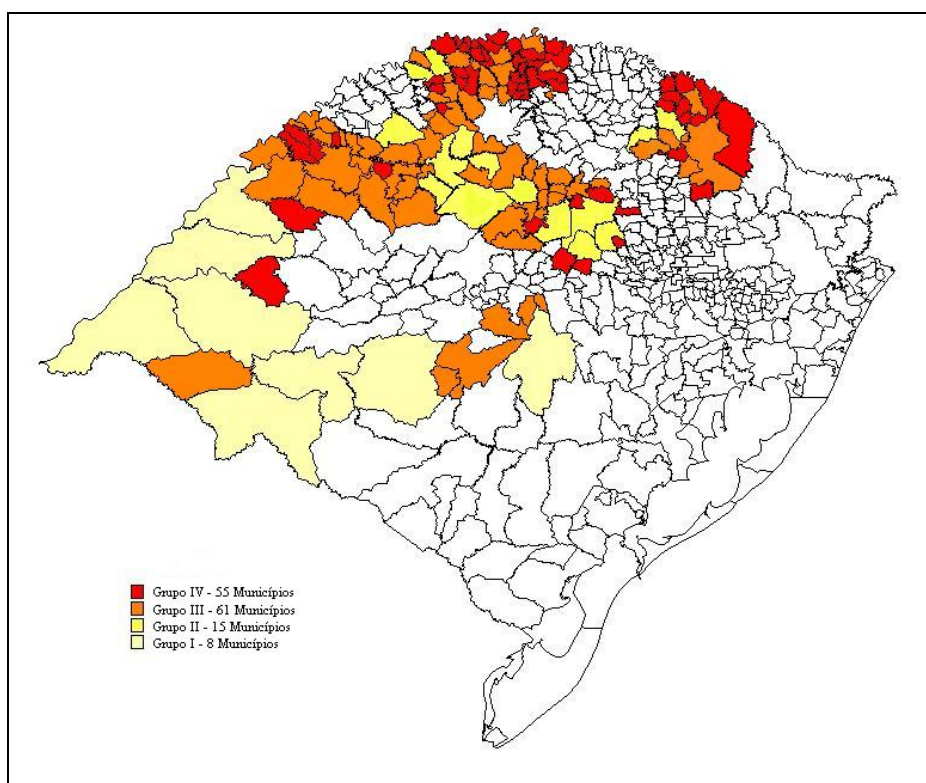
Os municípios fazem parte dos COREDEs essencialmente agrícolas, conforme classificação da Fundação de Economia e Estatística (FEE) junto ao Governo do Estado do Rio Grande do Sul (2005) são: a) Fronteira Oeste, com 10 municípios; b) Jacuí Centro, com 6 municípios; c) Missões, 21 municípios; d) Noroeste Colonial, com 29 municípios; e) Médio Alto Uruguai, também, com 29 municípios, f) Nordeste, com 18 municípios, g) Alto da Serra do Botucarái, com 14 municípios e, h) Alto Jacuí, com 12 municípios. O que pode ser verificado na ilustração.



Fonte: FEE.

ANEXO B – Mancha de Modernização dos Municípios que pertencem aos COREDEs essencialmente agrícolas

Os municípios fazem parte dos COREDEs essencialmente agrícolas, conforme classificação da Fundação de Economia e Estatística (FEE) junto ao Governo do Estado do Rio Grande do Sul (2005) segundo divisão política de 1993/94. Pois, ao período posterior diversos distritos se emanciparam e, dessa forma, os dados do novo Censo terão um acréscimo de municípios dos referidos COREDEs.



Fonte: Elaboração própria.