

Universidade de São Paulo  
Escola Superior de Agricultura “Luiz de Queiroz”

O Biplot na análise fatorial multivariada

Paula Cristina de Oliveira Klefens

Dissertação apresentada para obtenção do título de  
Mestre em Ciências. Área de concentração: Es-  
tatística e Experimentação Agronômica

Piracicaba

2009

Paula Cristina de Oliveira Klefens

Matemática

## O Biplot na análise fatorial multivariada

Orientador:

Prof. Dr. **CARLOS TADEU DOS SANTOS DIAS**

Dissertação apresentada para obtenção do título de Mestre em Ciências. Área de concentração: Estatística e Experimentação Agronômica

**Piracicaba**

**2009**

**Dados Internacionais de Catalogação na Publicação  
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Klefens, Paula Cristina de Oliveira

O Biplot na análise fatorial multivariada / Paula Cristina de Oliveira Klefens. - - Piracicaba, 2009.

102 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2009.  
Bibliografia.

1. Análise fatorial 2. Análise multivariada 3. SAS (Software Estatístico) I. Título

CDD 519.5354  
K63b

**"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"**

## Dedicatória

Aos meus pais Osvaldo e Susete,  
por tudo que sou, por todo amor e carinho  
dedicados à mim e por serem pessoas exemplares.

Aos meus irmãos Ricardo e Susiane, ao  
meu marido Gabriel e a minha amiga  
Paula Fonte por todo apoio.



## AGRADECIMENTOS

À **Deus** pela força para continuar.



**SUMÁRIO**

RESUMO . . . . .	9
ABSTRACT . . . . .	11
LISTA DE FIGURAS . . . . .	13
LISTA DE TABELAS . . . . .	15
1 INTRODUÇÃO . . . . .	17
1.1 Análise de Componentes Principais . . . . .	18
1.2 Análise Fatorial . . . . .	19
1.3 Análise Biplot . . . . .	20
2 REVISÃO DE LITERATURA . . . . .	23
2.1 Análise de Componentes Principais . . . . .	23
2.2 Análise Fatorial . . . . .	28
2.3 Biplot . . . . .	40
3 MATERIAL E MÉTODO . . . . .	45
3.1 Material . . . . .	45
3.2 Método de Pesquisa . . . . .	45
4 RESULTADOS E DISCUSSÃO . . . . .	49
5 CONSIDERAÇÕES FINAIS . . . . .	63
5.1 Conclusão . . . . .	63
5.2 Trabalhos Futuros . . . . .	63
REFERÊNCIAS . . . . .	65
ANEXOS . . . . .	67





## RESUMO

### O Biplot na análise fatorial multivariada

A análise multivariada é um conjunto de técnicas que são adequadas para situações onde várias variáveis correlacionadas estão envolvidas. Dentre essas técnicas temos as componentes principais e a análise fatorial. A técnica dos componentes principais reduz a dimensão de uma matriz de dados originais através de combinações lineares facilitando a interpretação desses dados e a análise fatorial que é o nome dado a uma classe de métodos estatísticos paramétricos (e não paramétricos) multivariados que correspondem a um grande número de métodos e técnicas que utilizam simultaneamente todas as variáveis do conjunto na interpretação do inter-relacionamento das variáveis observadas (COSTA, 2006). O objetivo da análise fatorial é descrever as relações de covariância entre algumas variáveis em algum termo subjacente, mas não observável, de quantidades aleatórias chamadas fatores (JOHNSON e WICHERN, 1998). Biplot é um gráfico estatístico, desenvolvido por Gabriel (1971), que representa no mesmo gráfico as variáveis e as observações com o intuito de demonstrar graficamente as relações existentes entre variáveis, entre observações e entre variáveis e observações. O presente trabalho tem como objetivo inserir a metodologia de análise biplot tridimensional na técnica de análise fatorial multivariada. Foi usado o software SAS para a realização da análise fatorial e a construção do gráfico biplot e um conjunto de dados para a aplicação do mesmo. O estudo mostra que o gráfico Biplot é um método de análise multivariada de suma importância quando inserido na análise fatorial facilitando e complementando a interpretação dos resultados.

Palavras-chave: Análise Fatorial; Multivariadas; Fatores; Biplot



## ABSTRACT

### The Biplot in multivariate factor analysis

Multivariate analysis is a set of techniques that are appropriate for situations where several correlated variables are involved. Among these techniques have the principal components and factor analysis. The technique of principal components reduces the size of an array of original data through linear combinations facilitating the interpretation of these data and factor analysis that is the name given to a class of parametric statistical methods (non parametric) multivariate corresponding to a large number of methods and techniques that use simultaneously all the variables set in the interpretation of the interrelationship of the observed variables (COSTA, 2006). The goal of factor analysis is to describe the covariance relationships between variables in any term underlying, but unobservable, random quantities called factors (Johnson and Wichern, 1998). Biplot is a statistical graph, developed by Gabriel (1971), which represents the in same graph the variables to the observations in order to demonstrate graphically the relationship between variables, between observations and between variables and observations. This paper aims to insert the biplot analysis methodology in three-dimensional technique of multivariate factor analysis. Was used SAS software to perform the factor analysis and construction of the biplot graph and a set of data for the application. The study shows that the graph Biplot is a method of analysis of the utmost importance when inserted in the factor analysis, facilitating and complementing the interpretation of results.

Keywords: Factor Analysis; Multivariate; Factors; Biplot



**LISTA DE FIGURAS**

Figura 1 - Rotação dos eixos no sentido da variabilidade máxima . . . . .	24
Figura 2 - Ortogonalidade dos fatores . . . . .	50
Figura 3 - Biplot da correlação entre variáveis e o fator 1 . . . . .	52
Figura 4 - Biplot da correlação entre variáveis e o fator 2 . . . . .	52
Figura 5 - Biplot da correlação entre variáveis e o fator 3 . . . . .	53
Figura 6 - Biplot das variáveis com correlação positiva . . . . .	54
Figura 7 - Biplot das variáveis com correlação negativa . . . . .	55
Figura 8 - Biplot das variáveis com correlação nula . . . . .	56
Figura 9 - Biplot das associações dos municípios do estado de Goiás com os fatores . . .	58
Figura 10 -Biplot das associações dos municípios de Tocantins com os fatores . . . . .	59
Figura 11 -Biplot da associação entre as variáveis e os objetos com destaque para as associações não obtidas na análise fatorial . . . . .	60



**LISTA DE TABELAS**

Tabela 1 - Os cinco primeiros autovalores, comunalidade dos primeiros fatores e porção de explicação da variabilidade dos dados . . . . .	49
Tabela 2 - Correlações entre os três fatores . . . . .	50
Tabela 3 - Valor das cargas fatoriais dos três fatores referentes as 24 variáveis indicadoras de modernização da pecuária após a rotação varimax . . . . .	51
Tabela 4 - Comunalidade dos primeiros fatores e porção de explicação da variabilidade dos dados após a rotação varimax . . . . .	53
Tabela 5 - Valores dos escores fatoriais dos três fatores referentes aos 26 municípios do estado de Goiás e Tocantins . . . . .	57





## 1 INTRODUÇÃO

Em muitas pesquisas científicas, se faz necessário o estudo do comportamento simultâneo de um conjunto de dados pelo fato da importância da influência exercida entre as variáveis. Essa influência pode ser analisada usando técnicas estatísticas multivariada em que podemos tirar conclusões valiosas sobre o quanto uma variável interfere em outra, ou o quanto uma variável interfere em um objeto, ou ainda o quanto um objeto interfere em outro objeto. Apesar de as técnicas multivariadas proporcionarem uma análise mais detalhada quando o estudo envolve várias variáveis, muitas vezes o resultado dessa análise é distorcido simplesmente por haver um grande número de variáveis, tornando difícil a interpretação da mesma.

Para resolver esse tipo de problema em que um grande número de variáveis pode atrapalhar a interpretação do resultado, existem, entre outras, duas técnicas que simplificam esse problema, a análise de componentes principais (APC) e a análise fatorial (AF). Estas análises envolvem técnicas que são adequadas para situações onde um grande número de variáveis correlacionadas está envolvido.

A análise de componentes principais tem como objetivo reduzir o número de variáveis de um conjunto de dados e explicar a estrutura de variâncias-covariâncias por poucas combinações lineares das variáveis originais. Já a análise fatorial tem como objetivo representar ou descrever um conjunto de variáveis originais a partir de um número menor de variáveis subjacentes, formando um conjunto menor de variáveis obtidas das originais. A interpretação desta última, é feita por meio de cargas fatoriais que são parâmetros do modelo e que expressam as covariâncias entre cada fator e as variáveis originais. No caso de utilizar variáveis padronizadas (matriz de correlações), esses valores correspondem às correlações entre os fatores e as variáveis originais (ARTES, 1998)

A variável observada é representada por uma função linear de uma pequena quantidade de fatores comuns, não observáveis, e uma variável latente específica. As covariâncias entre as variáveis observadas são geradas pelos fatores comuns e os termos específicos. Este último contribui somente para as variâncias de suas variáveis respostas relacionadas.

Muitas vezes, o resultado que obtemos nessas análises multivariadas não é su-

ficiente para a visualização do problema ou da solução. Pode-se afirmar que em muitos casos a visualização gráfica de um resultado torna-se mais fácil a conclusão a ser tirada. Mas aí é que entra um grande problema, como representar graficamente um conjunto multivariado?

Já existe na literatura um método gráfico para dados multivariados conhecido como Biplot. Com esse método é possível representar as  $p$  variáveis existentes em um conjunto de dados em um único gráfico bi ou tridimensional. O único problema é que, dentre esses métodos de redução de dados, até hoje o gráfico biplot só foi usado o método de componentes principais, que é um método puramente matemático e que tem como objetivo a redução na dimensão da matriz de dados originais e não no estudo da correlação e inter-correlação dos dados.

## 1.1 Análise de Componentes Principais

Algumas técnicas estatísticas multivariadas envolvem todas as relações entre os conjuntos de variáveis. No entanto, a análise de componentes principais (PCA) e a análise fatorial (AF) concentram relacionamentos dentro de um único conjunto de variáveis. Ambas as técnicas são usadas para reduzir a dimensionalidade do conjunto de variáveis, ou seja, para descrever os indivíduos ou objetos em termos de um número menor de variáveis não correlacionadas com pouca perda de informação tanto quanto possível, embora mantendo o máximo da informação da variação presente no conjunto de dados originais.

Pode-se dizer, assim como Moroco (2003, p.231), que “a análise de componentes principais é uma técnica de análise exploratória multivariada que transforma um conjunto de variáveis correlacionadas num conjunto menor de variáveis não correlacionadas, combinações lineares das variáveis originais”.

Para Johnson e Wichern (1998) apud Moroco (2003), algebricamente, as componentes principais são combinações lineares de  $p$  variáveis originais populacionais correlacionadas que tem como objetivo formar  $p$  combinações lineares independentes em que os pesos sejam estimados de tal forma que a primeira componente principal explique a maior proporção da variância nas variáveis originais, a componente seguinte explique a maior proporção da variância não explicada pela primeira componente que é independente da primeira e que a

escala da nova componente seja fixa de modo a manter constante a variância total.

Ferreira (1996) afirma que componentes principais é uma técnica de análise intermediária, pois esse método se presta fundamentalmente como um passo intermediário em análises estatísticas subsequentes, é um meio de simplificar dados pela redução do número de variáveis e não é baseada em um modelo estatístico e, portanto, não se constitui em um método final e conclusivo. Essa técnica pode ser aplicada nas técnicas de regressão múltipla, na análise de agrupamento e como estimadores de parâmetros nas técnicas multivariadas denominadas análise fatorial.

## 1.2 Análise Fatorial

Historicamente, relatado por Barroso e Artes (2003), a origem das técnicas de análise fatorial está ligada a estudo da área de psicologia onde Spearman em 1904, foi quem iniciou os estudos, desenvolvendo um método para a criação de um índice geral de inteligência (fator  $g$ ) com base nos resultados de vários testes que refletiam essa aptidão. Tratava de um primeiro método de análise fatorial adequado para a estimação de um único fator. Em seguida Lawley em 1940, deu maior rigor matemático em termos de inferência estatística, assumindo a distribuição normal para o vetor aleatório  $\mathbf{X}$ .

Para Härdle e Simar (2003) a análise fatorial e de componentes principais utilizam o mesmo conjunto de ferramentas matemáticas (decomposição espectral, projeções, rotações, etc). A princípio, pode-se tirar a conclusão de que essas análises possuem a mesma visão e estratégia produzindo resultados muito semelhantes. Mas a verdade é que existem grandes diferenças entre estas duas análises. A análise de componentes principais maximiza a variância projetada nos eixos e baseia-se na decomposição em valores singulares. Enquanto na análise fatorial, é o pós-rotação que fornece a máxima interpretação e envolve grandes variedades de procedimentos numéricos.

Já para Rencher (2001) a análise fatorial e a análise de componentes principais são semelhantes em vários aspectos, mas diferem em outros como, por exemplo, os componentes principais são definidas como combinações lineares das variáveis originais, isto é, os componentes são gerados em função das variáveis originais e uma grande parte da variância total dessas variáveis é explicada por poucos componentes. Já na análise fatorial, as variáveis

originais são expressas como combinações lineares de fatores não observáveis e busca-se explicar a covariância ou a correlação entre as variáveis.

A análise fatorial tem como objetivo descrever a ação de um conjunto grande de variáveis em termos de alguns fatores comuns, isto é, estuda uma quantidade de variáveis originais usando um número menor de fatores obtidos a partir dessas variáveis originais. Assume que cada variável original pode ser escrita como uma combinação linear desses fatores mais um termo residual (independente das variáveis).

### 1.3 Análise Biplot

Métodos práticos para executar os cálculos da análise multivariada têm sido desenvolvidos desde 1940, mas foram nos últimos 30 anos que os métodos se tornaram razoavelmente fáceis de serem executados pelos pesquisadores com o uso de um computador (MANLY, 2008).

Já existem muitos pacotes estatísticos e programas computacionais disponíveis para os cálculos multivariados. E tão importante quanto os programas para esses cálculos são os programas de representação gráfica para facilitar o entendimento dos resultados dessas análises multivariadas.

Mas como seria possível representar uma variável contra outras três ao mesmo tempo em alguma extensão da representação bi ou tridimensional? Então, o problema é ainda maior para mostrar, de uma maneira simples, o relacionamento que existe entre observações individuais em um conjunto multivariado de dados onde estas observações são descritas por quatro ou mais variáveis cada uma.

Para solucionar esse problema, Gabriel (1971), desenvolveu a técnica multivariada de representação gráfica - o Biplot, que é um gráfico de marcadores de dados de linha e coluna que formam uma tabela de dupla entrada. Os marcadores são calculados a partir da decomposição por valor singular da matriz de dados. O biplot é usado por alguns métodos multivariados para mostrar a relação existente entre variáveis, entre observações e entre variáveis e observações (LIPKOVICH e SMITH, 2002).

Até hoje o biplot é usado para obter a visualização gráfica de dois (ou três) primeiros componentes principais obtidas pela decomposição em valores singulares de uma

matriz de covariância amostral.

O objetivo do presente trabalho é dar um passo mais adiante, inserir uma metodologia de análise Biplot tridimensional na análise fatorial multivariada, representando as variáveis e observações em um único gráfico, permitindo a visualização das correlações e associações facilitando a interpretação dos resultados.



## 2 REVISÃO DE LITERATURA

### 2.1 Análise de Componentes Principais

A análise de componentes principais está relacionada com a explicação da estrutura de variâncias e covariâncias por meio de combinações lineares dos dados originais (FERREIRA, 1996). Os componentes amostrais dependem somente da matriz de covariâncias  $\mathbf{S}$  e não necessitam da normalidade multivariada dos dados. No entanto, para obter os escores dos componentes é preciso ter os dados originais (MANLY 2008).

Seus objetivos são a redução da dimensão original e obtenção de combinações lineares que interpretem as variáveis originais. A explicação da variabilidade total das  $p$  variáveis só é interpretada por  $p$  componentes principais, mas é possível explicar uma grande parte dessa variabilidade com um número  $k$  menor de componentes,  $k \leq p$ . Neste caso, os  $k$  componentes podem substituir as  $p$  variáveis originais e este conjunto é reduzido para um conjunto menor constituído de  $n$  medidas feitas em  $k$  componentes principais (JOHNSON e WICHERN, 1998).

Geometricamente, essas combinações lineares representam a seleção de um novo sistema de coordenadas obtida por rotação do sistema original  $X_1, X_2, \dots, X_p$  na direção de variabilidade máxima dos dados e fornecem uma descrição mais simples da estrutura de covariância.

Observe o exemplo bivariado, note que a variabilidade dos dados é maior no eixo  $\hat{X}_1$ :



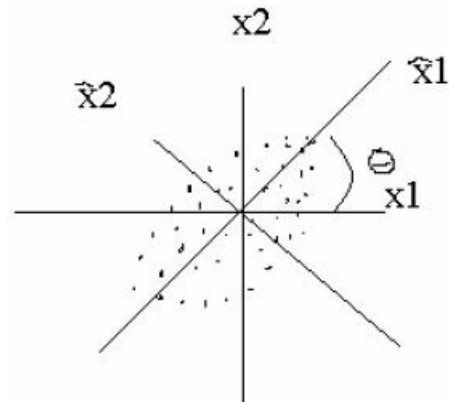


Figura 1 - Rotação dos eixos no sentido da variabilidade máxima

Componentes principais são as combinações lineares das variáveis mutuamente ortogonais que maximizam a variância total da amostra. Assim, seja a amostra aleatória  $X_1, X_2, \dots, X_p$  com vetor de média  $\bar{\mathbf{X}}$  e matriz de covariância amostral  $\mathbf{S}$  com autovalores  $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p \geq 0$ , retirada de uma população  $p$ -variada qualquer, com média  $\boldsymbol{\mu}$  e matriz de covariância  $\boldsymbol{\Sigma}$  com autovalores  $\lambda_1, \lambda_2, \dots, \lambda_p \geq 0$ .

Seja  $\mathbf{X}$  o vetor aleatório definido por:

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}$$

Considere a combinação linear tal que o  $i$ -ésimo componente principal de  $\hat{Y}$  é dado por:

$$\hat{Y}_i = \hat{\mathbf{e}}_i' \mathbf{X} = \hat{e}_{i1} X_1 + \hat{e}_{i2} X_2 + \dots + \hat{e}_{ip} X_p \quad (1)$$

em que  $\mathbf{e}$  é o vetor de autovetores de  $\mathbf{X}\mathbf{X}'$  e  $i = 1, \dots, p$ , então o modelo de CP estimado é dado por:

$$Y = \mathbf{e}' \mathbf{X} \quad (2)$$

O estimador da variância amostral é dado por:

$$\hat{V}ar(\hat{Y}_k) = \hat{V}ar(\hat{\mathbf{e}}'_k \hat{\mathbf{X}}) = \hat{\mathbf{e}}'_k \hat{V}ar(\hat{\mathbf{X}}) \hat{\mathbf{e}}_k = \hat{\mathbf{e}}'_k \mathbf{S} \hat{\mathbf{e}}_k$$

e

$$Cov(Y_i, Y_k) = cov(\hat{\mathbf{e}}'_i \hat{\mathbf{X}}, \hat{\mathbf{e}}'_k \hat{\mathbf{X}}) = \hat{\mathbf{e}}'_i \mathbf{S} \hat{\mathbf{e}}_k$$

Note que não há uma máximo para variância, pois quanto maior for  $\mathbf{e}$ , maior será o valor da variância. Dessa forma, seu máximo será obtido maximizando a forma quadrática

$$\lambda = \frac{\mathbf{e}' \mathbf{S} \mathbf{e}}{\mathbf{e}' \mathbf{e}} \quad (3)$$

obtendo  $(\mathbf{S} - \hat{\lambda} \mathbf{I}) \hat{\mathbf{e}} = \mathbf{0} \Leftrightarrow \mathbf{S} \mathbf{e} = \lambda \mathbf{e}$ , em que os autovalores são restritos ao comprimento unitário. A solução obtida por essa maximização conduz aos pares de autovalores e autovetores  $(\hat{\lambda}_i, \hat{\mathbf{e}}_i)$  de  $\mathbf{S}$ , que correspondem a variância amostral e a combinação linear que define os componentes principais.

Dessa forma obtém-se

$$\hat{V}ar(\hat{Y}_i) = \hat{\mathbf{e}}'_i \mathbf{S} \hat{\mathbf{e}}_i = \hat{\mathbf{e}}'_i \hat{\lambda}_i \hat{\mathbf{e}}_i = \hat{\lambda}_i \hat{\mathbf{e}}'_i \hat{\mathbf{e}}_i = \hat{\lambda}_i \quad (4)$$

e

$$Cov(Y_i, Y_k) = \hat{\mathbf{e}}'_i \mathbf{S} \hat{\mathbf{e}}_k = \hat{\mathbf{e}}'_i \hat{\lambda}_k \hat{\mathbf{e}}_k = \hat{\lambda}_k \hat{\mathbf{e}}'_i \hat{\mathbf{e}}_k = 0 \quad (5)$$

pela construção ortogonal dos autovetores ( $\mathbf{e}$ ). Note que

$$\sum_{i=1}^p Var(X_i) = \sum_{i=1}^p Var(Y_i) \Leftrightarrow \sum_{i=1}^p S_{ii} = \sum_{i=1}^p \lambda_i \quad (6)$$

Da decomposição espectral de  $\mathbf{S}$  dada por  $\mathbf{S} = \hat{\mathbf{P}} \hat{\Lambda} \hat{\mathbf{P}}'$  em que  $\hat{\mathbf{P}}$  é a matriz de autovetores de  $\mathbf{S}$ ,  $\hat{\mathbf{P}}'$  é a transposta de  $\hat{\mathbf{P}}$  e  $\hat{\Lambda}$  é a matriz diagonal dos autovalores de  $\mathbf{S}$ , sabendo que  $\hat{\mathbf{P}} \hat{\mathbf{P}}' = \hat{\mathbf{P}}' \hat{\mathbf{P}} = \mathbf{I}$  e usando as propriedades de traço, tem-se  $tr(\mathbf{S}) = tr(\hat{\mathbf{P}} \hat{\Lambda} \hat{\mathbf{P}}')$ , então

$$tr(\mathbf{S}) = \sum_{i=1}^p S_{ii} = tr(\hat{\mathbf{P}} \hat{\Lambda} \hat{\mathbf{P}}') = tr(\hat{\Lambda} \hat{\mathbf{P}} \hat{\mathbf{P}}') = tr(\hat{\Lambda}) = \sum_{i=1}^p \hat{\lambda}_i \quad (7)$$

A variância total amostral, pode ser explicada de 80 à 90% pelo primeiro, segundo ou terceiro componente principal e por esse motivo é que podemos substituir as  $p$  variáveis originais por esses  $k$  componentes sem muita perda de informação. A proporção da variância total devido ao  $i$ -ésimo componente principal é dada por:

$$\left(\frac{\hat{\lambda}_i}{\hat{\lambda}_1 + \dots + \hat{\lambda}_p}\right) \times 100 = \left(\frac{\hat{\lambda}_i}{\sum_{i=1}^p \hat{\lambda}_i}\right) \times 100 \quad (8)$$

O coeficiente de correlação entre os componentes  $\hat{Y}_i$  e as variáveis  $X_k$  é

$$r_{\hat{Y}_i, X_k} = \frac{\hat{e}_{ik} \sqrt{\hat{\lambda}_i}}{\sqrt{S_{kk}}}, \quad i = 1, \dots, p \quad (9)$$

se  $\hat{Y}_1 = \hat{e}'_1 \mathbf{X}$ ,  $\hat{Y}_2 = \hat{e}'_2 \mathbf{X}$ , ...,  $\hat{Y}_p = \hat{e}'_p \mathbf{X}$  são os primeiros componentes principais obtidos da matriz de covariância  $\mathbf{S}$ . Esses coeficientes medem somente a contribuição univariada de uma variável  $X$  a um componente  $\hat{Y}$  e não dependem do tamanho da medida das variáveis originais.

Os componentes principais podem ser definidos por componentes principais amostrais centrados na média amostral  $\bar{X}$ , isto é,

$$\hat{Y}_i = e'_i (\mathbf{X}_i - \bar{\mathbf{X}}), \quad i = 1, \dots, p$$

Geralmente os componentes principais variam com as transformações nas escalas. As variáveis com maior variabilidade podem vir a ser predominante na construção das componentes, mascarando eventualmente, a presença de variáveis com menor variabilidade. Neste caso, recomenda-se o uso de variáveis padronizadas cuja média é zero e variância é 1 (ARTES, 1998). Dessa forma,

$$\mathbf{Z}_j = \mathbf{D}^{-\frac{1}{2}} (X_j - \bar{X}), \quad j = 1, \dots, n$$

em que  $\mathbf{D}^{-\frac{1}{2}}$  é a matriz diagonal da inversa da raiz das variâncias,  $\mathbf{D}^{-\frac{1}{2}} = \text{diag}\left(\frac{1}{\sqrt{S_{11}}}, \dots, \frac{1}{\sqrt{S_{pp}}}\right)$

O estimador da covariância de  $\mathbf{Z}$  é

$$\hat{cov}(\mathbf{Z}) = \mathbf{D}^{-\frac{1}{2}} \hat{cov}(\mathbf{X}) \mathbf{D}^{-\frac{1}{2}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{S} \mathbf{D}^{-\frac{1}{2}} = \mathbf{R} \quad (10)$$

e com isso obtemos o  $i$ -ésimo componente principal, a porção da variabilidade total explicada do  $i$ -ésimo componente e a correlação amostral

$$\hat{Y}_i = \hat{e}'_i \mathbf{Z}, \quad i = 1, \dots, p \quad (11)$$

$$\text{variância explicada}(\hat{Y}_i) = \left(\frac{\hat{\lambda}_i}{p}\right) \times 100 \quad (12)$$

$$r_{Y_i, X_k} = \hat{e}_{ik} \sqrt{\hat{\lambda}_i} \quad (13)$$

A presença de dependência linear no conjunto de dados é indicada quando os últimos autovalores apresentam valores pequenos. Quanto ao número de componentes a serem escolhidos, deve-se considerar a quantidade da variação amostral explicada, o tamanho relativo dos autovalores e a interpretação subjetiva dos componentes (FERREIRA, 1996).

## 2.2 Análise Fatorial

Para estudos de dependências estruturais multinormais, as técnicas de explicações das covariâncias das respostas são as mais indicadas. As técnicas dos componentes principais também podem ser usadas para essa finalidade, mas não é a mais indicada, pois é uma transformação e não um resultado de um modelo fundamental da estrutura de covariância (FEREIRA, 1996).

Além da redução de dados, sumarização e análise das relações entre variáveis tentando explicá-las em termos de suas dimensões subjacentes comuns, a análise fatorial é uma técnica de interdependência, isto é, não há explicitada uma variável dependente (ALMEIDA, 2008).

Na análise fatorial o objetivo também é estudar a variação em uma quantidade de variáveis originais usando um número menor de fatores. Cada variável original pode ser expressa como uma combinação linear desses fatores, mais um termo residual que reflete o quanto uma variável é dependente da outra (FERREIRA, 1996).

Um ponto de semelhança entre as técnicas de dependência e a análise fatorial é dado onde cada variável observada é uma variável dependente, escrita em função de um conjunto de fatores obtidos a partir dessas variáveis observadas.

Para Barroso e Artes (2003, p.71) "a análise fatorial é uma técnica estatística que tem como objetivo descrever a estrutura de dependência de um conjunto de variáveis através dos fatores, que são variáveis que medem aspectos comuns e não são observadas".

Já Reis (1977) apud Barroso e Artes (2003), "define a análise fatorial como um conjunto de técnicas estatísticas cujo objetivo é representar ou descrever o conjunto de variáveis originais a partir de um número menor de variáveis hipotéticas". Isto é, formar um conjunto menor de variáveis a partir da estrutura de dependência existente entre as variáveis originais. A AF também permite saber o quanto cada fator está associado a cada variável e o quanto cada conjunto de fatores explica a variabilidade total dos dados originais.

A análise fatorial é uma técnica multivariada que estuda a correlação entre vetores aleatórios em termos de uma função linear dada por um número menor de variáveis aleatórias não observadas chamadas de fatores comuns, que geram as estruturas de covariância de  $X$ , mais uma variável única chamada de fator específico, que faz a variância dos termos

(GIRI, 1995).

As variáveis  $X_1, X_2, \dots, X_p$  geram como combinações lineares um número pequeno de variáveis aleatórias não observáveis  $F_1, F_2, \dots, F_m$  ( $m < p$ ) que são chamados fatores. O fator assemelha-se com as variáveis originais, pois ambas variam de indivíduo para indivíduo, porém diferem em não poder ser medidos ou observados (RENCHEER, 2001). Parte da variabilidade total dos dados é explicada pelos  $p$  fatores expressa através das somas das variâncias das variáveis (ARTES, 1998).

O modelo fatorial ortogonal pressupõe que as variáveis de um vetor  $\mathbf{X}_{p \times 1}$  sejam linearmente dependentes, expressas em termos de algumas variáveis aleatórias  $F_1, F_2, \dots, F_m$  denominadas *fatores comuns* e  $p$  fontes de variação  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ , denominadas *erros* ou *fatores específicos*. Em geral o modelo populacional de análise fatorial é:

$$X_i - \mu_i = \ell_{i1}F_1 + \ell_{i2}F_2 + \dots + \ell_{im}F_m + \varepsilon_i, \quad (14)$$

O coeficiente  $\ell_{ij}$  é chamado de *carga* da  $i$ -ésima variável no  $j$ -ésimo fator, então a matriz  $\mathbf{L}$  é a *matriz de cargas fatoriais*. Note que o  $i$ -ésimo fator específico  $\varepsilon_i$  está associado somente com a  $i$ -ésima variável  $X_i$  correspondente.

Matricialmente o modelo é:

$$\mathbf{X}_{p \times 1} - \boldsymbol{\mu}_{p \times 1} = \mathbf{L}_{p \times m} \mathbf{F}_{m \times 1} + \boldsymbol{\varepsilon}_{p \times 1} \quad (15)$$

o modelo acima e as suposições em termos dos operadores de esperanças e covariâncias a seguir, formam o modelo fatorial ortogonal:

$$E(\mathbf{F}) = \mathbf{0},$$

$$cov(\mathbf{F}) = E(\mathbf{F}\mathbf{F}') = \mathbf{I}$$

(isso equivale a dizer que os elementos do vetor aleatório  $\mathbf{F}$  são não correlacionados ou independentemente distribuídos)

$$E(\boldsymbol{\varepsilon}) = \mathbf{0},$$

$$cov(\boldsymbol{\varepsilon}) = E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Psi} = \begin{pmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \psi_p \end{pmatrix}$$

(o que equivale a dizer que os elementos de  $\boldsymbol{\varepsilon}$  são identicamente distribuídos) e assumindo que  $\mathbf{F}$  e  $\boldsymbol{\varepsilon}$  são independentes,

$$cov(\boldsymbol{\varepsilon}, \mathbf{F}) = E(\boldsymbol{\varepsilon}, \mathbf{F}') = \mathbf{0}$$

Com isso a estrutura de covariância do vetor aleatório  $\mathbf{X}$  é

$$cov(\mathbf{X}) = \boldsymbol{\Sigma} = E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'] \quad (16)$$

Note que substituindo  $\mathbf{X} - \boldsymbol{\mu}$  das suposições (16) no modelo (15) verifica-se que:

$$\begin{aligned} (\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})' &= (\mathbf{LF} + \boldsymbol{\varepsilon})(\mathbf{LF} + \boldsymbol{\varepsilon})' \\ &= (\mathbf{LF} + \boldsymbol{\varepsilon})((\mathbf{LF})' + (\boldsymbol{\varepsilon})') \\ &= (\mathbf{LF})(\mathbf{LF})' + \boldsymbol{\varepsilon}(\mathbf{LF})' + (\mathbf{LF})\boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \end{aligned}$$

então,

$$\begin{aligned} cov(\mathbf{X}) &= \boldsymbol{\Sigma} \\ &= E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'] \\ &= E[(\mathbf{LF})(\mathbf{LF})' + \boldsymbol{\varepsilon}(\mathbf{LF})' + (\mathbf{LF})\boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] \\ &= E[\mathbf{LFF}'\mathbf{L}' + \boldsymbol{\varepsilon}\mathbf{F}'\mathbf{L}' + \mathbf{LF}\boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] \\ &= \mathbf{LE}(\mathbf{FF}')\mathbf{L}' + E(\boldsymbol{\varepsilon}\mathbf{F}')\mathbf{L}' + \mathbf{LE}(\mathbf{F}\boldsymbol{\varepsilon}') + E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \\ &= \mathbf{LL}' + \boldsymbol{\Psi} \end{aligned}$$

ou seja,  $cov(\mathbf{X}) = \mathbf{LL}' + \boldsymbol{\Psi}$ , que é uma maneira alternativa de expressar o modelo fatorial, facilitando na estimação de seus parâmetros posteriormente.

Logo,

$$\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}' + \boldsymbol{\Psi} \quad (17)$$

Da mesma forma podemos encontrar as covariâncias entre os componentes

$$\begin{aligned} cov(\mathbf{X}, \mathbf{F}) &= E[(\mathbf{X} - \boldsymbol{\mu})\mathbf{F}'] \\ &= E[(\mathbf{L}\mathbf{F} + \boldsymbol{\varepsilon})\mathbf{F}'] \\ &= E[\mathbf{L}\mathbf{F}\mathbf{F}' + \boldsymbol{\varepsilon}\mathbf{F}'] \\ &= \mathbf{L}\mathbf{I} + \mathbf{0} \\ &= \mathbf{L} \end{aligned}$$

Então,

$cov(\mathbf{X}, \mathbf{F}) = \mathbf{L}$ , em que  $\mathbf{L}$  é a matriz de cargas fatoriais e  $cov(X_i, F_j) = \ell_{ij}$ , em que  $\ell_{ij}$  é o coeficiente chamado de carga da  $i$ -ésima variável no  $j$ -ésimo fator. Concluindo então que, as cargas fatoriais são as covariâncias medidas entre as variáveis observadas e os fatores comuns. Por outro lado, há uma relação entre as variâncias das variáveis originais e as cargas fatoriais como mostra a seguir:

$$\begin{aligned} Var(X_i) &= E(X_i^2) - [E(X_i)]^2 \\ &= E[(X_i - \mu_i)^2] - [E(X_i - \mu_i)]^2 \\ &= E[(\ell_i F_i + \varepsilon_i)^2] - [E(\ell_i F_i + \varepsilon_i)]^2 \\ &= \ell_i^2 + \psi_i \\ &= \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 \\ &= \sigma_{ii} \end{aligned}$$

$$\begin{aligned} cov(X_i, X_k) &= E(X_i - \mu_i)(X_k - \mu_k)' \\ &= \ell_{i1}\ell_{k1} + \ell_{i2}\ell_{k2} + \dots + \ell_{im}\ell_{km} \\ &= \sigma_{ik} \end{aligned}$$



em que  $\sigma_{ii}$  é a variância da  $i$ -ésima variável e  $\sigma_{ik}$  mostra o quanto a variável  $i$  está correlacionada (ou associada) com a variável  $j$ . Calculando a correlação existente entre o fator e cada variável observada, facilita a interpretação do fator comum

$$\begin{aligned} \text{corr}(X_i, F_j) &= \frac{\text{cov}(X_i, F_j)}{\sqrt{\text{Var}(X_i)\text{Var}(F_j)}} \\ &= \frac{\ell_{ij}}{\sigma_i} \end{aligned}$$

(JOHNSON e WICHERN, 1998).

A partir daí, podemos estimar o modelo Fatorial. Note que esse método de estimação usado será baseado no método das componentes principais (2), em que a matriz de variância amostral  $\mathbf{S}$  é descrita em termos dos pares de autovalores e autovetores  $(\hat{\lambda}_i, \hat{e}_i)$ . Esse nome, método de estimação por componentes principais, resulta do fato de que a carga dos coeficientes dos primeiros fatores são os componentes principais da amostra. Observe que podemos escrever o modelo de componentes da seguinte forma:

$$Y = (\mathbf{e}'\mathbf{\Lambda}^{-1/2})(\mathbf{\Lambda}^{1/2}\mathbf{X})$$

em que  $\mathbf{\Lambda}^{1/2}$  é a matriz diagonal de autovalores de  $\mathbf{X}$ . Considere

$\mathbf{e}'\mathbf{\Lambda}^{-1/2} = \mathbf{F} \Rightarrow$  Matriz de fatores comuns;

$\mathbf{\Lambda}^{1/2}\mathbf{X} = \mathbf{L}' \Rightarrow$  Matriz de cargas fatoriais.

Quando os últimos  $(p - m)$  autovalores são pequenos nós os desconsideraremos, utilizando apenas os  $m$  primeiros componentes principais. Essa rejeição dos últimos componentes nos gera um erro ( $\varepsilon$ ). Dessa forma, a estimativa do modelo Fatorial é dada por:

$$\mathbf{X} = \mathbf{L}'\mathbf{F} + \varepsilon \quad (18)$$

em que  $\mathbf{L}'\mathbf{F} + \varepsilon$  são desconhecidos. A matriz  $\mathbf{X}_{n \times p}$  é de observações,  $\mathbf{F}_{n \times m}$  é não observada contendo os  $m$  fatores comuns para as  $n$  observações,  $\mathbf{L}_{p \times m}$  é a matriz de cargas desconhecidas dos  $m$  fatores e  $\varepsilon_{n \times p}$  é a matriz não observada com os valores dos  $p$  fatores específicos para as  $n$  observações.

A vantagem de se usar estimação do modelo fatorial por componentes principais é o fato desse método não pressupor a normalidade dos dados envolvidos e por gerar estimativas simples e diretas dos escores fatoriais.

Quando a dependência de um conjunto de dados é analisada pode ocorrer casos em que as variáveis não são proporcionais, isto é, as variáveis possuem grandezas diferentes, então podemos encontrar dificuldades na interpretação dessa dependência. O método de componentes principais é um bom exemplo de método de estimação com grande sensibilidade às diferenças de grandezas. Nestes casos, é recomendável trabalhar com variáveis padronizadas

$$\mathbf{z}_j = \begin{pmatrix} \frac{(x_{j1} - \bar{x}_1)}{\sqrt{s_{11}}} \\ \frac{(x_{j2} - \bar{x}_2)}{\sqrt{s_{22}}} \\ \vdots \\ \frac{(x_{jp} - \bar{x}_p)}{\sqrt{s_{pp}}} \end{pmatrix}$$

cuja matriz de covariâncias  $\mathbf{S}$  é a matriz de correlação  $\mathbf{R}$  das observações  $x_1, x_2, \dots, x_n$ . A padronização evita o problema de ter uma variável com uma influência indevida na determinação das cargas fatoriais. Desse modo, é necessário algumas adaptações nos resultados obtidos anteriormente:

- A decomposição (17) deve ser feita sobre a matriz de correlações dos dados:

$$\boldsymbol{\rho} = \mathbf{L}\mathbf{L}' + \boldsymbol{\Psi}$$

- A correlação entre as variáveis originais e os fatores dão as cargas fatoriais:

$$\text{corr}(X_i, F_j) = \frac{\text{cov}(X_i, F_j)}{\sqrt{\text{Var}(X_i)\text{Var}(F_j)}} = \ell_{ij}$$

- As comunalidades passam a ser interpretadas como a proporção da variabilidade das variáveis padronizadas explicada pelos fatores:

$$\text{Var}(Z_i) = 1 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 + \psi_i$$

- A média das cargas ao quadrado é a proporção da variância total explicada pelo fator  $j$ , em que  $\sigma_T^2 = p$ . Note que trata-se da média das cargas fatoriais ao quadrado.

$$\lambda_j = \sum_{i=1}^p \frac{\ell_{ij}^2}{p}$$

- A média das comunalidades é a proporção da variabilidade total dos dados padronizados explicada pelo conjunto de fatores, em que  $\sigma_T^2 = p$ . Note que trata-se da média das comunalidades.

$$\frac{\sum_{i=1}^p c_i^2}{p}$$

A escolha do número de fatores é um passo muito importante na análise fatorial, pois se o pesquisador optar por reduzi-lo muito poderá ter dificuldades em identificar estruturas importantes existentes nos dados e se optar por um número excessivo de fatores talvez seja difícil de interpretá-los.

Observe que se o número de fatores for igual ao número de variáveis ( $m = p$ ), a matriz de variâncias e covariâncias pode ser reproduzida exatamente por  $\mathbf{LL}'$  ( $\mathbf{\Sigma} = \mathbf{LL}'$ ), ou seja, a matriz de variância específica é nula,  $\mathbf{\Psi} = \mathbf{0}$ , e a soma dos autovalores corresponde a soma das variâncias dessas variáveis e neste caso não há necessidade do uso da análise fatorial. Caso esteja sendo usada variáveis padronizadas (matriz de correlações) essa soma será igual ao número de variáveis  $p$  (ARTES, 1998).

Agora, se o número de fatores for menor que o número de variáveis ( $m < p$ ), a análise de fatores é ideal, pois o número de parâmetros  $p(m + 1)$  na análise de fatores, que são reproduzidos por  $pm$  pesos de fatores  $\ell_{ij}$  e  $p$  variâncias específicas  $\psi_i$ , são bem menores do que os  $p(p + 1)/2$  parâmetros de  $\mathbf{\Sigma}$ . Apesar de ideal, a dificuldade de fatorar  $\mathbf{\Sigma}$ , transforma a análise de fatores num grande problema. Quando  $m$  é muito menor que  $p$ , podemos obter soluções inconsistentes de acordo com interpretações estatísticas. Sendo  $m > 1$ , encontramos várias soluções consistentes que nos leva ao uso de uma importante ferramenta da análise fatorial, a rotação fatorial, que veremos mais adiante (JOHNSON e WICHERN, 1998).

Na literatura há vários critérios de escolha do número de fatores, mas neste trabalho trataremos do *critério da porcentagem da variância explicada*. Esse critério determina o número de fatores a ser escolhido de modo que o conjunto de fatores comuns explique

uma porcentagem pré-definida da variabilidade global. Normalmente essa porcentagem fica em torno de 80% da explicação da variância total amostral.

A porcentagem de explicação da variabilidade pode ser encontrada da seguinte forma:

$$\left( \begin{array}{l} \text{Porcentagem de} \\ \text{explicação da} \\ \text{variabilidade total} \\ \text{devido ao} \\ \text{\textit{j}-ésimo fator} \end{array} \right) = \left\{ \begin{array}{l} \frac{\hat{\lambda}_j}{\sum_{i=1}^p S_{ii}} \times 100, \quad \text{se AF for realizada em S;} \\ \frac{\hat{\lambda}_j}{p} \times 100, \quad \text{se AF for realizada em R.} \end{array} \right. \quad (19)$$

Determinado o número de fatores, podemos extrair informações sobre cargas fatoriais (que são parâmetros do modelo e expressam a covariância entre cada fator e as variáveis originais) e sobre a comunalidade (que é o quanto uma variável é explicada pelo número de fatores comuns escolhidos, denotada por  $h_i^2$ ).

Artes (1998) diz que as comunalidades são índices atribuídos às variáveis originais que expressam, em termos percentuais, o quanto da variabilidade de cada variável é explicada pelo modelo de AF estimado e isso significa que quanto mais próximo de 1 estiverem as comunalidades, melhor será o ajuste.

Já para Johnson e Wichern (1998), a *comunalidade* é a porção da variância da  $i$ -ésima variável explicada pelos  $m$  fatores comuns denotada por  $h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2$ . Para facilitar o entendimento, Barroso e Artes (2003) sugerem a utilização de  $\bar{h}_i^2 = \frac{h_i^2}{\sigma_i^2}$ , pois mantêm essa medida dentro do intervalo  $[0, 1]$  facilitando a interpretação da comunalidade. Esperam-se valores altos para  $h_i^2$  e então se tem um bom modelo.

A porção da variância da  $i$ -ésima variável ( $\sigma_{ii}$ ) devido ao fator específico é normalmente chamada de singularidade ou variância específica  $\Psi$  que é a contribuição do fator específico  $\psi_i = \sigma_{ii} - \sum_{j=1}^m \ell_{ij}^2$ . Assim,  $\sigma_{ii} = h_i^2 + \psi_i$ .

Considere  $\sum_{i=1}^m \sigma_{ii} = \sigma_T^2$  (a variabilidade total dos dados) e  $\sum_{i=1}^p h_i^2$  (a parcela de  $\sigma_T^2$  que é explicada pelo conjunto de fatores). Então temos que a parte da variabilidade total dos dados que é explicada pelo conjunto de fatores comuns é

$$\frac{\sum_{i=1}^p h_i^2}{\sigma_T^2}$$

As comunalidades são elementos da diagonal de  $\mathbf{LL}'$ . A proposta da análise de fatores é a determinação da matriz  $\mathbf{L}$  com os elementos da matriz de fator específico  $\mathbf{\Psi}$ , isto é, determina o valor das cargas relacionadas com os elementos específicos, tal que

$$\mathbf{\Sigma} - \mathbf{\Psi} = \mathbf{LL}' \quad (20)$$

Se a comunalidade total dos dados encontrada é alta, pode-se pensar que a variância específica foi bem pequena e então o erro ou resíduo  $\boldsymbol{\varepsilon} = \mathbf{\Sigma} - (\mathbf{LL}' + \mathbf{\Psi})$  foi suficientemente pequeno podendo ser ignorado, obtendo então  $\mathbf{\Sigma} = \mathbf{LL}'$  (GIRI, 1995). Esta situação nos apresenta uma similaridade com a análise de componentes principais, deixando claro que a AF, neste caso, não é interessante.

Mas a partir daí a análise fatorial pode fornecer um grande problema com a interpretabilidade dos fatores, pois é possível encontrar infinitos fatores que expliquem muito bem o comportamento dos dados para um mesmo conjunto de dados. Esse problema é geralmente resolvido introduzindo algumas limitações no modelo original. O procedimento que usaremos será o de *rotação fatorial*, que é uma técnica para girar os eixos de referência de dados, em torno da origem, até alcançar a posição ideal. Esta rotação pode ser ortogonal ou oblíqua.

Neste trabalho falaremos apenas da rotação fatorial *ortogonal*, que mantém os fatores independentes entre si, pois deixa pesos fatoriais altos associados a algumas variáveis para um fator e baixo para os outros, facilitando a leitura e definindo mais claramente os grupos de variáveis que fazem parte de um fator estudado (ALMEIDA, 2008). Essa rotação mantém as propriedades estatísticas inalteradas, mesmo alterando a matriz de cargas fatoriais.

Seja a matriz de cargas fatoriais  $\mathbf{L}_{p \times m}$  e a sua ordem para a rotação é igual ao número de autovalores significativos. Suponha que essa matriz  $\mathbf{L}$  é rotacionada pela matriz ortogonal  $\mathbf{T}_{m \times m}$  em que o modelo (15) também pode ser escrito como  $\mathbf{X} = (\mathbf{LT})(\mathbf{T}'\mathbf{F}) + \boldsymbol{\varepsilon} + \boldsymbol{\mu}$  produzindo uma nova matriz  $\mathbf{L}^*$  tal que

$$\mathbf{L}^* = \mathbf{LT}$$

Então  $\mathbf{T}$  é uma matriz de transformação e sua ortogonalidade  $\mathbf{T}\mathbf{T}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$  faz com que as comunalidades fiquem inalteradas:

$$\sum_{j=1}^m \ell_{ij}^{*2} = \sum_{j=1}^m \ell_{ij}^2 \Leftrightarrow h_i^{*2} = h_i^2 \quad (21)$$

assim como as especificidades (JOHNSON e WICHERN, 1998).

O método de rotação mais utilizado, que é para quando  $m \geq 2$ , é o método *varimax* que segundo Ferreira (1996) foi proposto por KAISER (1958, 1959), o qual maximiza a variância dos pesos dos fatores em cada fator (dentro de cada coluna de  $\mathbf{L}$ ) sujeito a restrição de ortogonalidade dos fatores, ou seja, que os fatores permaneçam não correlacionados entre si. Com esse método esperamos encontrar que, para cada termo, existam apenas alguns pesos fatoriais significativos e todos os outros sejam próximos de zero. Isto é, tenta dividir as variáveis automaticamente em subconjuntos disjuntos, cada um associado a um fator (HÄRDLE e HLÁVKA, 2007).

Esse método proposto por Kaiser (1958, 1959) apud Johnson e Wichern (1998), está relacionado com a soma das variâncias das cargas fatoriais quadráticas dentro de cada coluna da matriz de fatores  $\mathbf{L}$ . Por esse critério, procura-se por cargas fatoriais  $\ell_{ij}$  que minimizam

$$V^* = \frac{1}{p^2} \sum_{j=1}^m [p \sum_{i=1}^p \ell_{ij}^4 - (\sum_{i=1}^p \ell_{ij}^2)^2] \quad (22)$$

em que  $V^*$  é uma medida de variabilidade das cargas fatoriais. Com esse critério, respostas com comunalidades grandes ou pequenas passam a ter pesos iguais. A melhora desse critério sugerida por Kaiser é dada por:

$$V^* = \frac{1}{p^2} \sum_{j=1}^m [p \sum_{i=1}^p X_{ij}^4 - (\sum_{i=1}^p X_{ij}^2)^2] \quad (23)$$

em que

$$X_{ij}^4 = \frac{\ell_{ij}}{\sqrt{\sum_{j=1}^m \ell_{ij}^2}}$$

então o conjunto  $\ell_{ij}$  que tiver a variância máxima será escolhido pela a rotação (JOHNSON e WICHERN, 1998).

Além de analisar a estrutura de correlação encontrando os parâmetros do modelo, podemos encontrar também, o valor ou escores para cada fator, que são variáveis não observáveis, mas que podem ser estimadas. Essas estimativas dos fatores são os escores fatoriais que como não são parâmetros do modelo, são valores atribuídos às variáveis latentes, não podem ser estimados da maneira usual.

As formas mais usadas para se estimar os escores fatoriais são pelos métodos *dos mínimos quadrados ponderados* ou *da regressão*. Este trabalho focará apenas o método dos mínimos quadrados ponderados, pois este não pressupõe que fatores e erros sejam normalmente distribuídos assim como no método de regressão.

Segundo Johnson e Wichern (1998), o uso do método dos mínimos quadrados ponderados para encontrar os escores fatoriais foi sugerido por Bartlett. Mas apenas se  $\mu$ ,  $\mathbf{L}$ ,  $\Psi$  forem conhecidos do modelo (15) e  $Var(\varepsilon_i) = \psi_i$ . Esse método propõe escolher as estimativas  $\hat{f}$  de  $f$  para minimizar a soma de quadrados dos erros ponderados pela recíproca da variância

$$\sum_{i=1}^p \left( \frac{\varepsilon_i^2}{\psi_i} \right) = \varepsilon' \Psi \varepsilon = (\mathbf{x} - \boldsymbol{\mu} - \mathbf{L}\mathbf{f})' \Psi^{-1} (\mathbf{x} - \boldsymbol{\mu} - \mathbf{L}\mathbf{f}) \quad (24)$$

e usando as estimativas  $\hat{\mathbf{L}}$ ,  $\hat{\Psi}$  e  $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}$  como valores verdadeiros, obtemos o  $j$ -ésimo escore fatorial

$$\hat{\mathbf{f}}_j = (\hat{\mathbf{L}}' \hat{\Psi}^{-1} \hat{\mathbf{L}})^{-1} \hat{\mathbf{L}}' \hat{\Psi}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}})$$

ou seja, estimaremos um escore fatorial pontual com o mínimo de erro possível.

Como as cargas fatoriais foram estimadas pelo método dos componentes principais, não obteremos os escores fatoriais pela ponderação de mínimos quadrados, mas pelo método dos mínimos quadrados ordinais onde é suposto que  $\psi_i$  são iguais ou quase iguais. Dessa forma a equação dos escores fatoriais será dada por:

$$\hat{\mathbf{f}}_j = (\hat{\mathbf{L}}' \hat{\mathbf{L}})^{-1} \hat{\mathbf{L}}' (\mathbf{x}_j - \bar{\mathbf{x}}) \quad (25)$$

ou

$$\hat{\mathbf{f}}_j = (\hat{\mathbf{L}}_z' \hat{\mathbf{L}}_z)^{-1} \hat{\mathbf{L}}_z' \mathbf{z}_j$$

para dados padronizados. Desde que  $\mathbf{L} = [\sqrt{\hat{\lambda}_1} \mathbf{e}_1 \ \sqrt{\hat{\lambda}_2} \mathbf{e}_2 \ \dots \ \sqrt{\hat{\lambda}_m} \mathbf{e}_m]$ , temos então

$$\hat{\mathbf{f}}_j = \begin{pmatrix} \frac{1}{\sqrt{\hat{\lambda}_1}} \hat{\mathbf{e}}_1' (\mathbf{x}_j - \bar{\mathbf{x}}) \\ \frac{1}{\sqrt{\hat{\lambda}_2}} \hat{\mathbf{e}}_2' (\mathbf{x}_j - \bar{\mathbf{x}}) \\ \dots \\ \frac{1}{\sqrt{\hat{\lambda}_m}} \hat{\mathbf{e}}_m' (\mathbf{x}_j - \bar{\mathbf{x}}) \end{pmatrix} \quad (26)$$



### 2.3 Biplot

Em muitas áreas, os pesquisadores utilizam os primeiros e mais importantes componentes de uma análise para agrupar observações e itens de acordo com a representação em duas ou no máximo três dimensões. Um gráfico bidimensional pode ser obtido representando as informações das observações para duas variáveis em um diagrama, já para um gráfico biplot, a idéia é de acrescentar informações sobre várias variáveis em um único gráfico obtido pelos componentes principais.

Biplot é uma técnica multivariada proposta por Gabriel (1971), e tem como objetivo a representação gráfica aproximada de uma matriz  $X$  de dados em outra  $Y_{n \times p}$  de baixa dimensão, permitindo a visualização das relações e inter-relação entre linhas e colunas de  $Y$  em um único gráfico (DIAS, 2007). Independente da forma de fatoração, as aplicações do método Biplot dividem-se em dois: a representação de uma matriz de dados com a finalidade descritiva e de diagnóstico de modelos.

Qualquer matriz de posto dois e qualquer tipo de variável (contínua ou discreta) podem ser exibidos como um biplot bidimensional que é formado por um vetor de cada linha e de cada coluna de uma matriz de tal forma que qualquer elemento de  $\mathbf{Y}$  seja exatamente o produto interno dos vetores correspondentes à sua linha e à sua coluna.

Segundo Roux e Gardner (2005), Gabriel (1971) introduziu o biplot como uma exibição gráfica de linhas e colunas de uma matriz de dados em um único gráfico, tornando-se populares e sendo aplicados em muitas áreas. Ainda por Roux e Gardner (2005), a abordagem moderna de Gower e Hand (1996)<sup>1</sup> tornam a metodologia biplot mais acessível, além da teoria unificada que engloba os diferentes Biplots que resultam do tradicional método biplot de Gabriel, fornece uma visão de uma teoria biplot multivariada. Esta abordagem tem a vantagem de que muito da funcionalidade da metodologia biplot pode ser facilmente entendida por não estatísticos.

O bi do nome Biplot refere-se a duas informações contidas na matriz de dados. Nas linhas encontramos informações pertencentes às amostras ou unidades amostrais e nas colunas informações que pertencem às variáveis. O biplot permite observar a posição de uma observação com relação à outra observação e a associação de cada uma das variáveis

---

<sup>1</sup>GOWER, J. C. e HAND, D. J. (1996). Biplots

em relação a qualquer observação ou a outra variável. Assim, pode-se observar como as observações se agrupam e quais variáveis contribuem para sua posição dentro dessa representação (JOHNSON e WICHERN, 1998).

### Construindo biplot (componentes principais)

A construção de um biplot bidimensional origina-se dos componentes principais amostrais. Seja  $\mathbf{X}_{n \times p}$ . Procura-se por uma aproximação  $\mathbf{Y}_{n \times p}$  de posto 2 da matriz original  $\mathbf{X}$ , então obteremos um exato biplot bidimensional de  $\mathbf{Y}$ .

A melhor aproximação de posto 2 de  $\mathbf{X}$  é obtida decompondo-a em valor singular (DVS), que consiste em escrevê-la como produto de três matrizes:

$$\mathbf{X}_{n \times p} = \mathbf{U}_{n \times p} \mathbf{\Lambda}_{p \times p} \mathbf{V}'_{p \times p} \quad (27)$$

em que  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ , que são valores singulares que obedecem  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ ,

$\mathbf{U}$ : matriz ortogonal de autovetores de  $\mathbf{X}_{n \times p} \mathbf{X}'_{p \times n}$ ;

$\mathbf{V}$ : matriz ortogonal de autovetores de  $\mathbf{X}'_{p \times n} \mathbf{X}_{n \times p}$ .

(JOHNSON e WICHERN, 1998)

Assim, pela decomposição em valores singulares

$$\mathbf{X} = \lambda_1 \mathbf{u}_1 \mathbf{v}'_1 + \lambda_2 \mathbf{u}_2 \mathbf{v}'_2 + \dots + \lambda_p \mathbf{u}_p \mathbf{v}'_p$$

ou seja,

$$\mathbf{X} = \sum_{i=1}^p \lambda_i \mathbf{u}_i \mathbf{v}'_i$$

em que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ ,  $\mathbf{u}_i$  é o vetor de autovetores associados ao  $i$ -ésimo autovalor da matriz  $\mathbf{X}\mathbf{X}'$  com dimensão  $(n \times 1)$  e  $\mathbf{v}'_i$  é o vetor linha de autovetores associados ao  $i$ -ésimo autovalor da matriz  $\mathbf{X}'\mathbf{X}$  com dimensão  $(1 \times p)$ . Se considerarmos os 2 primeiros componentes, obteremos uma aproximação através de uma matriz de classificação igual a 2. Ou seja,

$$\mathbf{Y}_{n \times p} \approx \mathbf{Y}_{(2)} = \lambda_1 \mathbf{u}_1 \mathbf{v}'_1 + \lambda_2 \mathbf{u}_2 \mathbf{v}'_2 \quad (28)$$

ou

$$\mathbf{Y}_{n \times p} \approx \mathbf{Y}_{(2)} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \mathbf{v}'_1 \\ \mathbf{v}'_2 \end{pmatrix}$$

Para obter o biplot da matriz  $\mathbf{Y}$ , é preciso escrevê-la como o produto de duas matrizes  $\mathbf{GH}'$ , onde  $\mathbf{G}$  é uma matriz ( $n \times 2$ ) e  $\mathbf{H}$  é uma matriz ( $2 \times p$ ). Isso pode ser feito de várias formas, dependendo da fatoração da matriz  $\mathbf{Y}$  em produtos de matrizes, tal que:

$$\mathbf{Y}_{n \times p} \approx (\mathbf{U}\mathbf{\Lambda}^\alpha)(\mathbf{\Lambda}^{1-\alpha}\mathbf{V}') = \mathbf{GH}' \quad (29)$$

em que  $\alpha$  variando geralmente em 0, 1/2 e 1 determina a ênfase que será dada as linhas ou as colunas de  $\mathbf{Y}$ . Gabriel propôs 3 formas básicas de construção das matrizes  $\mathbf{G}$  e  $\mathbf{H}$

- Biplot SQRT, em que nenhuma ênfase é dada a linha ou coluna de  $\mathbf{Y}$ ,  $\alpha = 1/2$ .

$$\mathbf{Y}_{n \times p} = \begin{pmatrix} u_{11}\sqrt{\lambda_1} & u_{21}\sqrt{\lambda_2} \\ \vdots & \vdots \\ u_{n1}\sqrt{\lambda_1} & u_{n2}\sqrt{\lambda_2} \end{pmatrix} \begin{pmatrix} v_{11}\sqrt{\lambda_1} \dots & v_{1p}\sqrt{\lambda_1} \\ v_{21}\sqrt{\lambda_2} \dots & v_{2p}\sqrt{\lambda_2} \end{pmatrix}$$

- Biplot  $JK$ , em que a ênfase é nas linhas de  $\mathbf{Y}$ ,  $\alpha = 1$ .

$$\mathbf{Y}_{n \times p} = \begin{pmatrix} u_{11}\lambda_1 & u_{21}\lambda_2 \\ \vdots & \vdots \\ u_{n1}\lambda_1 & u_{n2}\lambda_2 \end{pmatrix} \begin{pmatrix} v_{11} & \dots & v_{1p} \\ v_{21} & \dots & v_{2p} \end{pmatrix}$$

- Biplot  $GH$ , em que a ênfase é nas colunas de  $\mathbf{Y}$ ,  $\alpha = 0$ .

$$\mathbf{Y}_{n \times p} = \begin{pmatrix} u_{11} & u_{21} \\ \vdots & \vdots \\ u_{n1} & u_{n2} \end{pmatrix} \begin{pmatrix} v_{11}\lambda_1 & \dots & v_{1p}\lambda_1 \\ v_{1n}\lambda_2 & \dots & v_{2p}\lambda_2 \end{pmatrix}$$

Considere a  $i$ -ésima linha de  $\mathbf{G}$ ,  $\mathbf{g}'_i$  e a  $j$ -ésima coluna de  $\mathbf{H}'$ ,  $\mathbf{h}_j$

$$\mathbf{Y}_{n \times p} = \begin{pmatrix} \mathbf{g}'_1 \\ \vdots \\ \mathbf{g}'_n \end{pmatrix} \begin{pmatrix} \mathbf{h}_1 & \dots & \mathbf{h}_p \end{pmatrix}$$

O biplot consiste em plotar os vetores  $\mathbf{g}'_i$ ,  $i = 1, \dots, n$  e  $\mathbf{h}_j$ ,  $j = 1, \dots, p$  o espaço tridimensional. Eles são vetores que partem da origem até o ponto de coordenadas dadas pelos elementos de  $\mathbf{g}_i$  e  $\mathbf{h}_j$ .

Cada elemento  $y_{ij}$  de  $\mathbf{Y}$  é representado como produto interno de  $\mathbf{g}'_i \mathbf{h}_j$ . O comprimento da projeção de  $\mathbf{g}'_i$  sobre  $\mathbf{h}_j$  é dado por:

$$L_{p_{\mathbf{g}'_i/\mathbf{h}_j}} = \frac{|\mathbf{g}'_i \mathbf{h}_j|}{L_{\mathbf{h}_j}} \Rightarrow |\mathbf{g}'_i \mathbf{h}_j| = \begin{cases} L_{\mathbf{h}_j} L_{p_{\mathbf{g}'_i/\mathbf{h}_j}} \\ \text{ou} \\ L_{\mathbf{g}'_i} L_{p_{\mathbf{g}'_i/\mathbf{h}_j}} \end{cases} = L_{\mathbf{g}'_i} L_{\mathbf{h}_j} \cos(\theta) \quad (30)$$

O produto interno  $\mathbf{g}'_i \mathbf{h}_j$  pode ser visto como o produto do comprimento de um dos vetores vezes o comprimento da projeção do outro no primeiro (DIAS, 2008).



### 3 MATERIAL E MÉTODO

#### 3.1 Material

Os dados utilizados no presente estudo foram retirados do trabalho de Dias (1991). Estes se referem ao mapeamento de 26 municípios de Goiás e Tocantins (Anexo B). De cada município foram medidas 24 variáveis indicadoras de modernização da pecuária no processo tecnológico (Anexo A). Essas variáveis encontram-se listadas no anexo A e os municípios no anexo B.

Foram utilizados do software SAS os comandos PROC FACTOR como ferramenta de execução para a obtenção da análise fatorial usual e o PROC IML como ferramenta de construção da matriz de escores fatoriais e do gráfico Biplot para fatores.

#### 3.2 Método de Pesquisa

O método de pesquisa estatística utilizado no desenvolvimento deste trabalho foi a análise fatorial multivariada expressa como:

$$\mathbf{X}_{p \times 1} - \boldsymbol{\mu}_{p \times 1} = \mathbf{L}_{p \times m} \mathbf{F}_{m \times 1} + \boldsymbol{\varepsilon}_{p \times 1}$$

em que

$\mathbf{X}$  : vetor de dados originais;

$\boldsymbol{\mu}$  : vetor de média populacional;

$\mathbf{L}$  : matriz de cargas fatoriais;

$\mathbf{F}$  : matriz de fatores comuns;

$\boldsymbol{\varepsilon}$  : vetor de fatores específicos (ou erros).

Como vimos, a *matriz de cargas fatoriais* fornece o valor dos parâmetros do modelo o que permite observar o quanto cada variável e cada autovalor está associado a cada fator. A *matriz de fatores comuns* que são variáveis latentes obtida pela combinação linear das variáveis originais, explicam parte da variabilidade total dos dados através da soma das variâncias das variáveis, cada fator pode representar estruturas subjacentes dos dados e, também, mostram o quanto cada variável é dependente da outra. E a *matriz de fatores específicos* que está associada somente a uma única variável.

Foi usado o procedimento de rotação fatorial ortogonal para resolver o problema de inconsistência das respostas pelo método varimax proposto por Kaiser (1958, 1959) apud Johnson e Wichern (1998).

O número de fatores foi obtido pelo método das considerações a priori e não pelo critério da porcentagem da variância explicada, como descrito na revisão de literatura, simplesmente para manter o objetivo do trabalho.

Até o momento, identificamos variáveis importantes analisando a matriz de cargas fatoriais e para fins de análises posteriores obtivemos a matriz de escores fatoriais, que nos fornece a estrutura de correlação e a influência dos fatores sobre as observações. Essa primeira parte da análise com o programa de análise fatorial encontra-se em anexo (Anexo C).

Considere  $\mathbf{X}_{n \times p}$  uma matriz de dados originais. Queremos construir o gráfico Biplot tridimensional dessa matriz de dados. Para isso precisamos de uma matriz  $\mathbf{Y}_{n \times p}$  que seja a aproximação de posto 3 de  $\mathbf{X}$ . Decompondo  $\mathbf{X}$  por valores singulares (DVS) obteremos a melhor aproximação de posto 3. Assim temos,

$$\begin{aligned}\mathbf{X}_{(n \times p)} &= \mathbf{U}_{(n \times p)} \mathbf{\Lambda}_{(p \times p)} \mathbf{V}'_{(p \times p)} \\ &= \lambda_1 \mathbf{u}_1 \mathbf{v}'_1 + \lambda_2 \mathbf{u}_2 \mathbf{v}'_2 + \dots + \lambda_p \mathbf{u}_p \mathbf{v}'_p\end{aligned}$$

em que

$\mathbf{U}$  : matriz de autovetores de  $\mathbf{X}_{(n \times p)} \mathbf{X}'_{(p \times n)}$

$\mathbf{V}$  : matriz de autovetores de  $\mathbf{X}'_{(p \times n)} \mathbf{X}_{(n \times p)}$

$\mathbf{\Lambda}$  : matriz  $\text{diag}(\lambda_1, \dots, \lambda_p)$

Considerando os 3 primeiros componentes principais teremos uma aproximação através de uma matriz de classificação igual a 3. Ou seja,

$$\mathbf{Y}_{n \times p} \approx \mathbf{Y}_{(3)} = \lambda_1 \mathbf{u}_1 \mathbf{v}'_1 + \lambda_2 \mathbf{u}_2 \mathbf{v}'_2 + \lambda_3 \mathbf{u}_3 \mathbf{v}'_3 \quad (31)$$

Escrevendo  $\mathbf{Y}$  como produto de 2 matrizes  $\mathbf{GH}'$ , obtemos o Biplot da mesma, em que  $\mathbf{G}$  é uma matriz  $n \times r$  e  $\mathbf{H}$  uma matriz  $r \times p$ . Em particular, obteremos apenas o

biplot SQRT, em que nenhuma ênfase é dada a linha e a coluna.

$$\mathbf{Y} = \begin{pmatrix} u_{11}\sqrt{\lambda_1} & u_{12}\sqrt{\lambda_2} & u_{13}\sqrt{\lambda_3} \\ u_{21}\sqrt{\lambda_1} & u_{22}\sqrt{\lambda_2} & u_{23}\sqrt{\lambda_3} \\ \vdots & \vdots & \vdots \\ u_{n1}\sqrt{\lambda_1} & u_{n2}\sqrt{\lambda_2} & u_{n3}\sqrt{\lambda_3} \end{pmatrix} \begin{pmatrix} v_{11}\sqrt{\lambda_1} & v_{12}\sqrt{\lambda_1} & \dots & v_{1p}\sqrt{\lambda_1} \\ v_{21}\sqrt{\lambda_2} & v_{22}\sqrt{\lambda_2} & \dots & v_{2p}\sqrt{\lambda_2} \\ v_{31}\sqrt{\lambda_3} & v_{32}\sqrt{\lambda_3} & \dots & v_{3p}\sqrt{\lambda_3} \end{pmatrix}$$

Com este estudo mostraremos que é possível colocar em um único gráfico todas as informações de correlações e inter-correlações contidas em uma matriz de dados, facilitando a interpretação tanto dos dados obtidos por essa análise quanto à posição do mesmo no gráfico. O programa para a obtenção do gráfico biplot tridimensional encontra-se no anexo D.





## 4 RESULTADOS E DISCUSSÃO

Pelo resultado da análise fatorial realizada para os dados que se encontram no anexo A obtida pelo software SAS, observou-se que os cinco primeiros autovalores são maiores que um e que quatro deles explicam 80,9% da variância total das 24 variáveis. Isso mostra que existem quatro conjuntos de variáveis originais correlacionadas entre si e pouco ou não correlacionadas entre eles, o que nos dá a idéia de um conjunto de dados relativamente heterogêneo.

Pela porcentagem de explicação da variabilidade total das 24 variáveis (Tabela 1) pode-se observar que os três fatores retirados para a análise explicam 73,91% da variância total e em particular é o primeiro fator que explica a maior parte (47,93%). Note que apesar dessa porcentagem de explicação estar abaixo dos 80% como descrito na revisão de literatura, desenvolvemos o estudo com apenas três fatores para manter o objetivo do trabalho, pois o intuito é construir o gráfico Biplot com um número pequeno de fatores retirando o máximo de informação possível e também por haver uma certa dificuldade em visualizar um gráfico quadridimensional.

Tabela 1 - Os cinco primeiros autovalores, comunalidade dos primeiros fatores e porção de explicação da variabilidade dos dados

	1	2	3	4	5	
Autovalores	11,5029724	4,396811	1,8384187	1,6849472	1,3852929	% de explicação
Comunalidade dos 5 fatores	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	
	0,4793	0,1832	0,0766	0,070206	0,057720	80,90%

As cargas fatoriais antes da rotação Varimax (Anexo E) mostram que o primeiro fator ( $F1$ ) está forte e positivamente associado com as variáveis  $V8$  (0,88),  $V5$  (0,87),  $V18$  (0,86),  $V13$  (0,85),  $V9$  (0,84),  $V7$  (0,83),  $V12$  (0,82),  $V21$  (0,80) e  $V10$  (0,71) e, forte e negativamente associado com as variáveis  $V23$  (-0,84)  $V6$  (-0,82) e  $V20$  (-0,80). O segundo fator ( $F2$ ) possui alta associação e positiva somente com a variável  $V2$  (0,83) e muito forte e negativamente associação com a variável  $V11$  (-0,92). Já o terceiro fator ( $F3$ ) possui somente associação negativa com a variável  $V19$  (-0,4).

Durante a análise, utilizando o software SAS, houve a necessidade de padronizar os dados. Então, os resultados que nos interessa é após a padronização e a rotação varimax a qual mantém a ortogonalidade dos dados (Tabela 2 e figura 2).

Tabela 2 - Correlações entre os três fatores

Fatores	Fator 1	Fator 2	Fator 3
Fator 1	1,0000	0,0000	0,0000
Fator 2	0,0000	1,0000	0,0000
Fator 3	0,0000	0,0000	1,0000

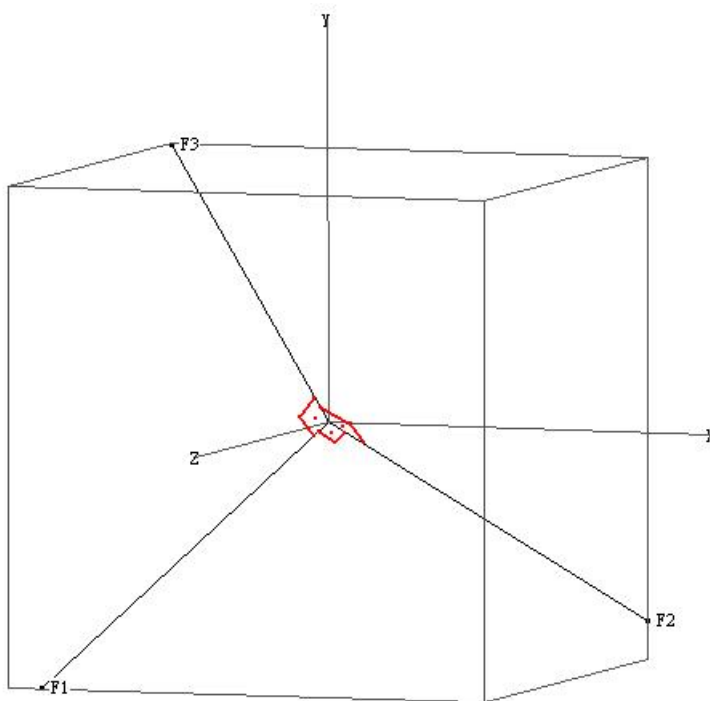


Figura 2 - Ortogonalidade dos fatores

Pela tabela 3 pode-se observar que o primeiro fator ( $F1$ ) passou a ter associação forte e positiva com as variáveis  $V21$  (0, 91),  $V22$  (0, 86),  $V17$  (0, 87) e  $V18$  (0, 79) e, forte, mas negativamente correlacionada com as variáveis  $V23$  e  $V20$  (-0, 95). O segundo fator ( $F2$ ) está forte e positivamente associado com as variáveis  $V2$  (0, 92),  $V12$  (0, 88),  $V14$  (0, 85),  $V8$  (0, 76)

e  $V13$  (0, 72) e correlacionada negativamente apenas com a variável  $V11$  ( $-0,73$ ). E o terceiro fator  $F3$  possui correlação apenas positiva com as variáveis  $V3$  (0, 73) e  $V1$  (0, 70). Com esses resultados pode-se dizer, em resumo, que  $F1$  está representando a pecuária capitalista,  $F2$  a pecuária intensiva e  $F3$  a mecanização.

Tabela 3 - Valor das cargas fatoriais dos três fatores referentes as 24 variáveis indicadoras de modernização da pecuária após a rotação varimax

Variáveis	Fator 1	Fator 2	Fator 3	Variáveis	Fator 1	Fator 2	Fator 3
V1	-0,48725	-0,14283	<u>0,70223</u>	V13	0,52350	<u>0,72339</u>	-0,04445
V2	-0,23832	<u>0,91585</u>	-0,06950	V14	<u>0,10536</u>	0,85426	-0,09786
V3	-0,41063	-0,5549	<u>0,3064</u>	V15	0,38241	0,60412	-0,14956
V4	0,06982	-0,47318	-0,52106	V16	0,59478	0,23881	-0,18252
V5	0,63646	0,55482	-0,25273	V17	<u>0,86793</u>	-0,00904	0,00921
V6	-0,59458	-0,53932	0,20201	V18	<u>0,78661</u>	0,45443	0,21039
V7	0,54977	0,66212	-0,04804	V19	-0,47343	0,00533	-0,65805
V8	0,53507	<u>0,76542</u>	-0,00569	V20	<u>-0,94746</u>	-0,06975	0,06120
V9	0,62723	0,57012	-0,04069	V21	<u>0,91525</u>	0,08189	-0,16526
V10	0,59894	0,34735	-0,19482	V22	<u>0,86268</u>	0,00446	-0,04859
V11	0,52534	<u>-0,73252</u>	-0,17945	V23	<u>-0,94754</u>	-0,11029	0,15856
V12	0,38025	<u>0,88256</u>	0,05606	V24	0,05163	0,67975	-0,02718

Note que as cargas fatoriais nos fornece apenas o peso de influência das variáveis sobre os fatores, a real correlação entre estes pode ser vista de forma clara no gráfico Biplot ilustrado nas figuras 3, 4 e 5.

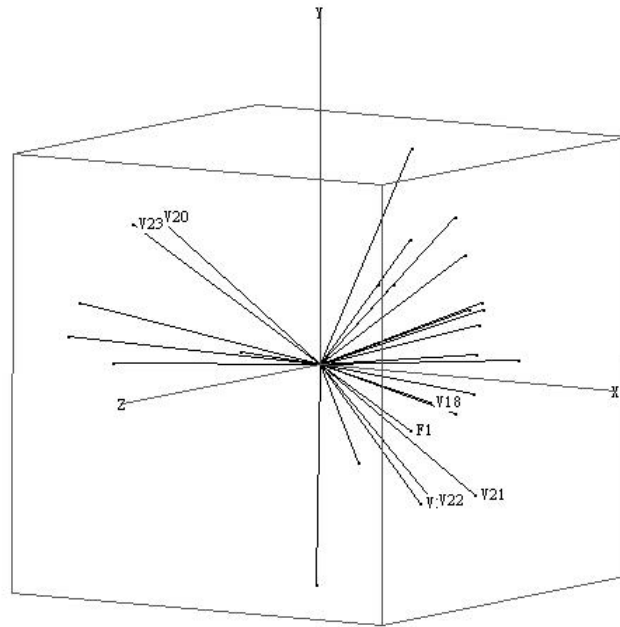


Figura 3 - Biplot da correlação entre variáveis e o fator 1

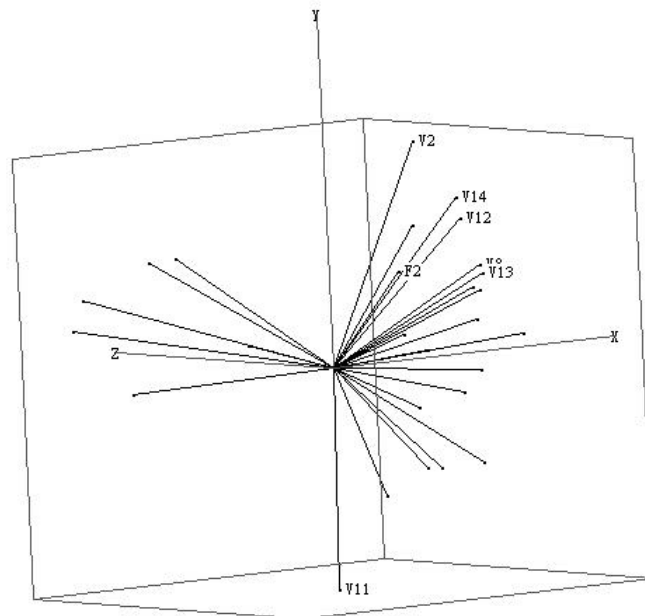


Figura 4 - Biplot da correlação entre variáveis e o fator 2

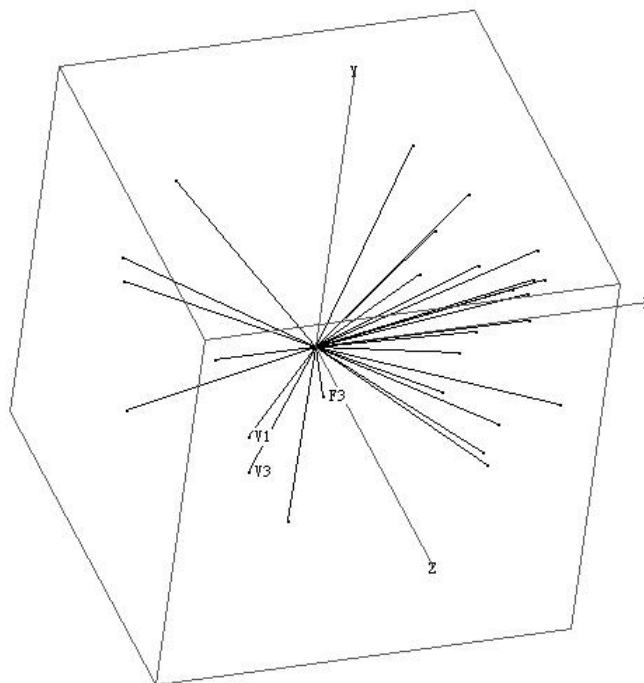


Figura 5 - Biplot da correlação entre variáveis e o fator 3

Como era de se esperar, após a rotação a porção de explicação da variabilidade de  $F1$  manteve-se maior que  $F2$  e  $F3$ , porém menor que a obtida anteriormente à rotação obtendo  $F1 : 36,45\%$ ,  $F2 : 28,75\%$  e  $F3 : 8,71\%$ , como pode ser visto na tabela 4. Isso mostra a importância desse método de rotação dos eixos pois os fatores ficaram mais claramente definidos em termos de associação entre eles e as variáveis.

Tabela 4 - Comunalidade dos primeiros fatores e porção de explicação da variabilidade dos dados após a rotação varimax

	Fator 1	Fator 2	Fator 3	% de explicação
Comunalidade dos 3 fatores	0,3645	0,2875	0,0871	73,91%

Note, no anexo F, que as variáveis positivas e altamente correlacionadas entre si são  $V1$  com  $V3$  (0,98),  $V20$  com  $V23$  (0,97),  $V8$  com  $V13$  (0,93),  $V8$  com  $V12$  (0,90),  $V5$  com

$V_{10}$  (0, 89),  $V_5$  com  $V_{13}$  (0, 88),  $V_{12}$  com  $V_{13}$  (0, 86),  $V_5$  com  $V_8$  e  $V_9$  com  $V_{15}$  (0, 85),  $V_2$  com  $V_{14}$  (0, 81) e  $V_{12}$  com  $V_7$  (0, 80). Já as variáveis  $V_{21}$  com  $V_{23}$  e  $V_{20}$  com  $V_{21}$  (-0, 92),  $V_{22}$  com  $V_{23}$  (-0, 89) e  $V_{20}$  com  $V_{22}$  (-0, 87) também estão fortemente correlacionadas, mas de forma negativa. Essa correlação entre as variáveis obtida pela análise fatorial fica mais clara quando vista graficamente e é exatamente isso que uma parte do Biplot faz tornar uma análise numérica em uma análise gráfica facilitando a interpretação dos resultados (Figura 6 e 7).

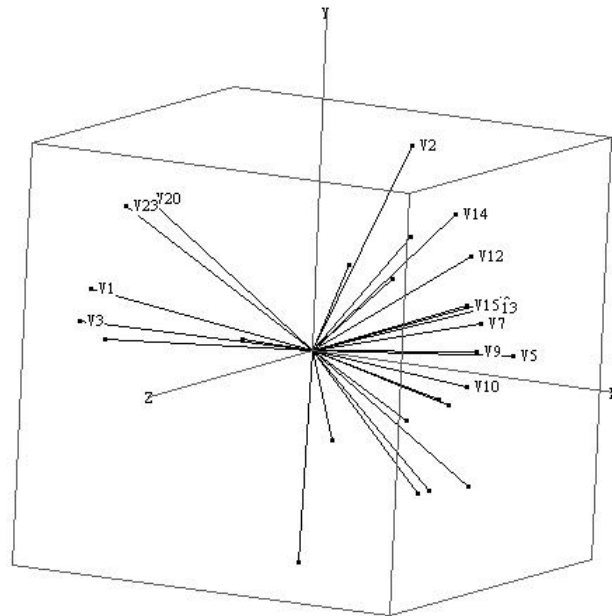


Figura 6 - Biplot das variáveis com correlação positiva

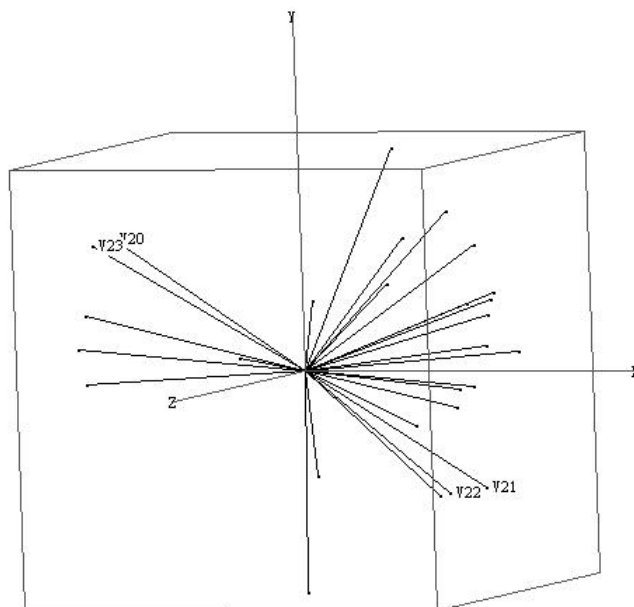


Figura 7 - Biplot das variáveis com correlação negativa

Chamamos de correlação nula as variáveis que não possuem covariabilidade e isso acontece quando a correlação entre elas é zero ou próximo de zero. Essa correlação nula é identificada visualmente quando duas variáveis formam ângulo de 90 graus (Figura 8).



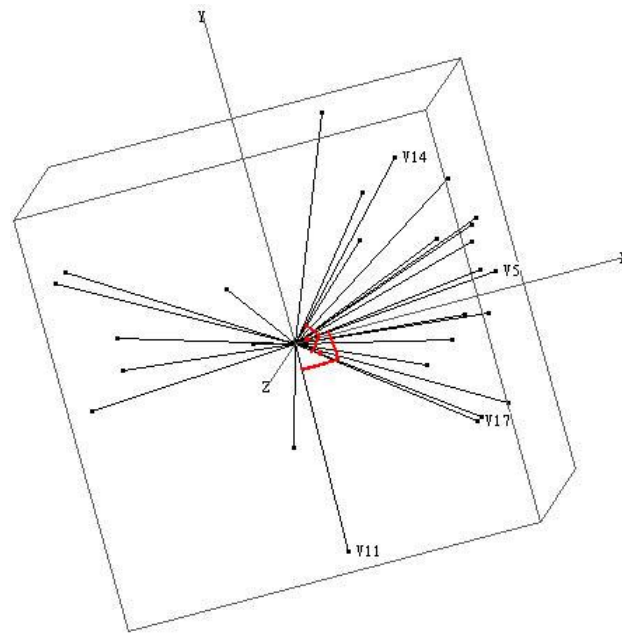


Figura 8 - Biplot das variáveis com correlação nula

Essa análise visual das correlações com os fatores também pode ser feita para as observações que no presente caso são os municípios. Essa análise ajuda a identificar que variáveis, representadas pelos fatores, estão associadas a qual ou quais observações. Na tabela 5 temos os valores encontrados para essa associação pela análise fatorial, chamados de escores fatoriais. Os municípios representados por *O1* até *O8* pertencem ao estado de Tocantins e os municípios representados por *O9* até *O26* pertencem ao estado de Goiás.

Tabela 5 - Valores dos escores fatoriais dos três fatores referentes aos 26 municípios do estado de Goiás e Tocantins

Observações (Municípios)	Fator 1	Fator 2	Fator 3	Observações (Municípios)	Fator 1	Fator 2	Fator 3
O1	<u>-1,73</u>	0,03	-0,08	O14	-0,81	<u>1,42</u>	-0,29
O2	-0,49	-0,52	-0,71	O15	-0,37	<u>2,57</u>	-0,08
O3	-0,64	-0,77	-0,35	O16	0,71	0,26	0,46
O4	0,15	-0,96	-0,50	O17	-0,09	<u>1,61</u>	-0,03
O5	-0,27	-0,77	-0,62	O18	0,54	1,06	0,70
O6	-0,88	-0,64	-0,46	O19	-0,93	-0,52	-1,02
O7	<u>-1,29</u>	-0,63	<u>3,74</u>	O20	0,02	-0,18	-0,30
O8	<u>-1,13</u>	-0,85	0,18	O21	<u>1,97</u>	-0,92	-0,52
O9	0,80	<u>-1,26</u>	-1,14	O22	0,82	<u>1,29</u>	-0,50
O10	0,74	-0,68	-0,62	O23	<u>1,16</u>	<u>1,09</u>	-0,16
O11	<u>1,33</u>	-0,76	2,14	O24	-0,15	0,81	-0,01
O12	<u>-1,22</u>	0,05	-0,59	O25	0,13	0,125	0,17
O13	-0,47	<u>-1,12</u>	-0,01	O26	<u>2,11</u>	0,01	0,60

A tabela 5 mostra que o estado de Goiás apresenta pecuária pouco capitalizada, pouco intensiva e mostra um baixo grau de modernização, pois dos dezoito municípios selecionados para análise são apresentadas apenas quatro associação forte e positiva e uma associação forte e negativa com  $F1$  que são  $O26$  (2, 11),  $O21$  (1, 97),  $O11$  (1, 33),  $O23$  (1, 16) e  $O12$  (-1, 22), cinco associações forte e positiva e duas negativas com  $F2$  que são  $O15$  (2, 57),  $O17$  (1, 61),  $O14$  (1, 42),  $O22$  (1, 29),  $O23$  (1, 09),  $O9$  (-1, 26) e  $O13$  (-1, 12). E para  $F3$ , apenas uma associações forte e positiva e duas negativas,  $O11$  (2, 14) e  $O9$  (-1, 14) e  $O19$  (-1, 02) respectivamente. Observe esses resultados no gráfico Biplot na figura 9.

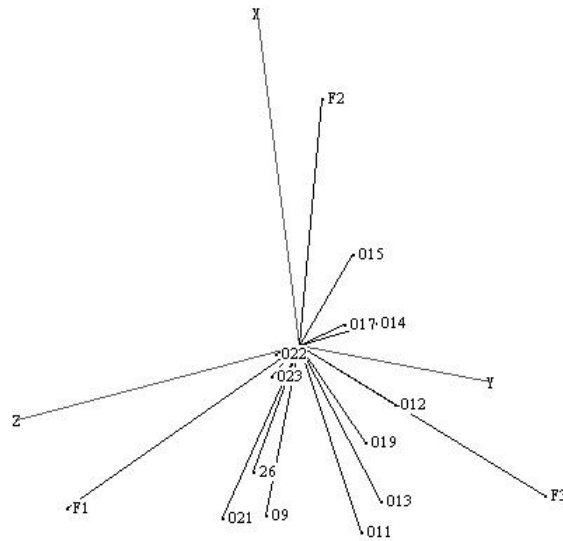


Figura 9 - Biplot das associações dos municípios do estado de Goiás com os fatores

Para o estado de Tocantins o quadro é ainda pior, pois dos oito municípios selecionados não foi observada associação positiva relevante para  $F1$  e três negativa para o mesmo fator  $O1$   $(-1,73)$ ,  $O7$   $(-1,29)$  e  $O8$   $(-1,13)$ . Nenhuma associação positiva ou negativa para  $F2$  que seja considerável para a análise e apenas uma associação relevante para  $F3$ , sendo ela positiva e muito fortemente associada  $O7$   $(3,74)$  levando a concluir que o estado de Tocantins possui uma pecuária com baixa capitalização, muito pouco intensiva e baixíssimo grau de modernização (Figura 10).

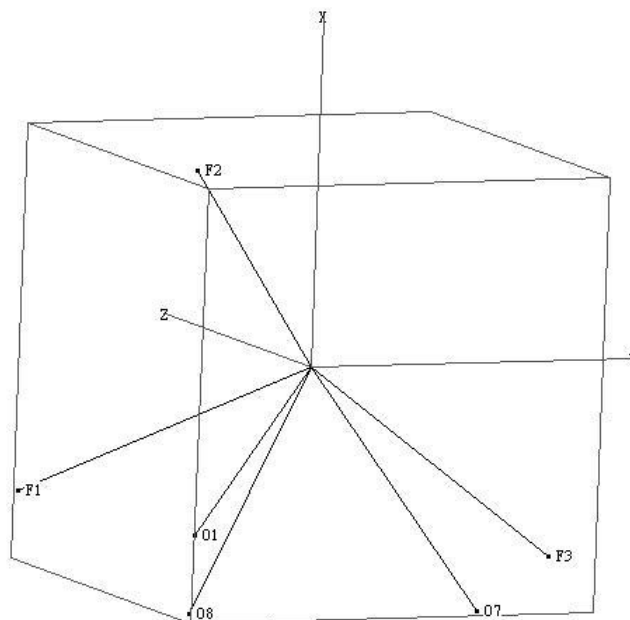


Figura 10 - Biplot das associações dos municípios de Tocantins com os fatores

Além do complemento à análise fatorial, o Biplot permite tirar algumas conclusões que outra análise não oferece. Foi possível ver que observações se associam as variáveis, mas por intermédio dos fatores, que representam um conjunto de variáveis que se associam a algumas observações. É importante aqui imaginar o quão mais fácil e interessante seria obter essas conclusões analisando e visualizando exatamente as variáveis que estiverem associadas as observações. Através da representação gráfica do Biplot, é possível obter essa informação que a análise fatorial não fornece.

Quando há a correlação de variáveis com um fator pode-se dizer que esse fator representa essas variáveis que estão correlacionadas com ele. Este quando associado com algumas observações, está representando as variáveis que estariam associadas a essas observações. Baseado nisso e nos resultados obtidos acima, podemos dizer que há uma certa associação das variáveis  $V_{21}$ ,  $V_{22}$ ,  $V_{17}$ ,  $V_{18}$ ,  $V_{23}$  e  $V_{20}$  com as observações  $O_{26}$ ,  $O_{21}$ ,  $O_{11}$ ,  $O_{23}$  e  $O_{12}$  obtidos pelo fator 1 na análise fatorial, das variáveis  $V_2$  e  $V_{11}$  com as observações  $O_{15}$ ,  $O_{17}$ ,  $O_{14}$ ,  $O_{22}$ ,  $O_{23}$ ,  $O_9$  e  $O_{13}$  obtidos pelo fator 2 e a variável  $V_{19}$  com as observações  $O_{11}$ ,  $O_9$  e  $O_{19}$  obtidos pelo fator 3, porém não se sabe que variável se associa a qual observação.

Neste ponto é que entra o diferencial da análise pelo gráfico Biplot, pois além de correlacionar variáveis com fatores e observações com fatores, obtemos também as associações entre variáveis e observações e variáveis, observações e fatores num único gráfico e permitindo por meio de vetores detectar quais variáveis estão associadas as quais observações.

Pela figura 11 observou-se que as variáveis *V14*, *V20* e *V24* estão associadas com as observações *O24*, *O1* e *O17* respectivamente, mas não aparecem nas associações obtidas pela análise fatorial usual. Veja que esse é outro ponto importante na análise pelo gráfico Biplot, conseguimos obter associações, muitas vezes cruciais para um resultado satisfatório, mas que não são exibidas nos resultados da análise fatorial pela perda de informação devido ao uso de fatores na representação das associações.

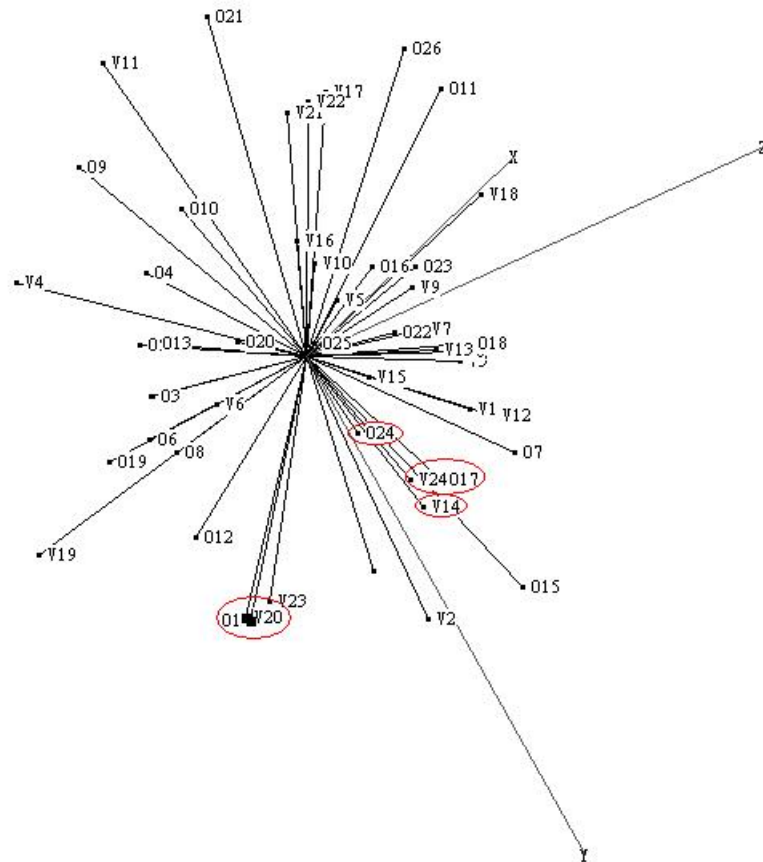


Figura 11 - Biplot da associação entre as variáveis e os objetos com destaque para as associações não obtidas na análise fatorial

Com essas associações obtidas analisando o gráfico Biplot podemos dizer que o município de Pires do Rio-GO tem alto valor de leite alimentar/área ocupada do estabelecimento, Bico do Papagaio-TO tem uma porcentagem alta do número de estabelecimento sem pessoal contratado e Anicuns-GO tem grande porcentagem de trabalho de empreitada/pessoal ocupado, mostrando que a análise Biplot tem um "insight" maior no entendimento e na interpretação das relações entre variáveis e observações do que a análise fatorial usual.



## 5 CONSIDERAÇÕES FINAIS

### 5.1 Conclusão

De acordo com o objetivo deste estudo e os resultados obtidos, pode-se concluir que:

- O gráfico Biplot é um método de análise multivariada que por si só revela resultados satisfatórios para qualquer conjunto de dados multivariado. Este aliado a outro método de análise, neste caso a análise fatorial, tem um papel importantíssimo como complemento para a interpretação dos resultados numéricos.
- Especificamente à análise fatorial, o método Biplot traz maneiras surpreendentes e de fácil visualização, no entendimento e na interpretação dos resultados, como a representação gráfica das variáveis e dos fatores conjuntamente (cargas fatoriais), mostrando cada subconjunto de variáveis representado por cada fator. E também a representação gráfica das observações e dos fatores (escores fatoriais).
- Além de colocar em um único gráfico todas as informações obtidas pela análise fatorial, o biplot mostra resultados que a mesma não fornece como associar variáveis com observações e associar variáveis, observações e fatores num único gráfico possibilitando identificar quais variáveis estão associadas a quais observações e se essas associações são positivas ou não.
- Além do benefício de mostra as associações entre variáveis e observações deixando claro qual variável está associada a qual observação, o Biplot permite identificar algumas associações que a análise fatorial não detecta pelo fato de usar uma única variável (fator) para representar um conjunto de variáveis.

Em resumo o método de análise Biplot é um ótimo aliado para a análise fatorial como complemento bem com, e talvez o mais importante, na conclusão final da análise.

### 5.2 Trabalhos Futuros

Como continuidade deste estudo, seria interessante que trabalhos futuros tratassem dos seguintes assuntos:



- Para facilitar a execução e melhorar a aparência do programa, sugiro que seja criada uma rotina no programa SAS para a obtenção dos escores fatoriais pelo método dos mínimos quadrados ponderados;
- Para uma análise fatorial mais elaborada seria interessante desenvolver uma metodologia de análise para fatore pelo método triplot.

## REFERÊNCIAS

- ALMEIDA, T. L. Departamento de Matemática. **Arquivos de aulas**. FURG. Disponível em: <<http://www.dmat.furg.br/ taba/posanafat.htm>>. Acesso em: 08 jul. 2008.
- ARTES, R. Aspectos Estatísticos da análise Fatorial de Escala de Avaliação. **Revista de Psiquiatria Clínica**. São Paulo: v.25, n.5, p.223-228, 1998.
- BARROSO, L. P.; ARTES, R. Análise Multivariada. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRÔNOMICA 10.; E REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA 48, 2003, Lavras. **Anais...** São Paulo: IBMEC, 2003. p. 71 - 103.
- COSTA, G. G. O. **Um procedimento inferencial para Análise Fatorial Utilizando as Técnicas Bootstrap e Jackknife**: Construção de Intervalos de Confiança e Testes de Hipóteses. 2006. p.22-40 Tese (Doutor em Engenharia Elétrica) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006.
- DIAS, C. T. S. Intranet do Departamento de Ciências Exatas. **Arquivos de aulas**. LCE - ESALQ - USP. Disponível em: <<http://www.lce.esalq.usp.br/tadeu/intranet/aula12.pdf>>. Acesso em: 29 set. 2008.
- DIAS, C. T. S. Representação gráfica bidimensional de uma matriz de dados multivariados: Biplot. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRÔNOMICA 12.; E REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA 52., 2007, Santa Maria. **Anais...** Santa Maria, 52a. rbras, 2007
- DIAS, C. T. S. **Estudo da modernização da Agropecuária dos estados de Goiás e Tocantins, pela Análise Fatorial**. Piracicaba: ESALQ, 1991. 21p.
- FERREIRA, D. F. **Análise Multivariada**. Lavras: UFLA, 1996. p.233-282 e p.309-354
- GABRIEL, K. R. The Biplot graphic display of matrices with application to principal component analysis. **Biometrika**, 58: p.453-467.
- GIRI, N. C. **Multivariate Statistical analysis**, Factor Analysis. New York: Marcel Dekker, 1995. 378p.
- HÄRDLE, W.; HLÁVKA, Z. **Multivariate Statistics: Exercises and Solutions**. Local: Springer, 2007. p.185-204
- HÄRDLE, W.; SIMAR, L. **Applied Multivariate Statistical Analysis**. Local: Tech, 2003. p.275-298.
- HARRIS, R. J. A. **Primer of Multivariate Statistics**. 3.ed. Mahwah: Lawrence Erlbaum Associates, 2001. p.394-433
- JOHNSON, R. A.; WICHERN, D.W. **Applied Multivariate Statistical Analysis**. 4.ed.

Upper Saddle River: Prentice Hall, 1998. p.458-586.

LIPKOVICH, I.; SMITH, E. P. Biplot and Singular Value Decomposition Macros for Excel. **Journal of Statistical Software**. Blacksburg, v.7, n.5, p.1 - 15, jun. 2002.

MANLY, B. J. F. **Métodos Estatísticos Multivariados**: uma introdução. 3.ed. Porto Alegre: Bookman, 2008. p.89-117.

MOROCO, J. Análise de Componentes Principais. In: ----- . **Análise Estatística com Utilização do SPSS**. 3 ed. Lisboa: Sílabo, 2003. cap. 9, p.230 - 257.

RENCHEER, A. C. **Methods of Multivariate Analysis**. Factor Analysis, 2.ed. : Wiley-Interscience, 2001. p.408-448

ROUX, N. J. LE; GARDNER, S. Analysing Your Multivariate Data as a Pictorial: A Case for Applying Biplot Methodology?. **International Statistical Review**. University of Stellenbosch, v.3, n.73, p.365–387, jun. 2005

## ANEXOS



**ANEXO A - As 24 variáveis retirada de cada município**

- V1: pessoal ocupado/número de tratores;
- V2: pessoal ocupado/1.000 ha de área ocupada;
- V3: área (1.000 ha) ocupada/número de tratores;
- V4: pessoal ocupado/ número de estabelecimento;
- V5: porcentagem de áreas trabalhada;
- V6: porcentagem de área produtiva não trabalhada mais lavoura temporária em descanso;
- V7: porcentagem de estabelecimento com adubação;
- V8: número de bovinos/ha de pastagem nativa mais porcentagem plantada;
- V9: porcentagem de estabelecimentos com conservação de solo;
- V10: porcentagem de pastagem plantada;
- V11: área relativa aos estabelecimentos;
- V12: produção de leite/área ocupada do estabelecimento;
- V13: valor bovinos/área ocupada do estabelecimento;
- V14: valor do leite alimentar/área ocupada do estabelecimento;
- V15: valor do leite industrial/área ocupada do estabelecimento;
- V16: valor da produção vegetal/pessoa ocupada;
- V17: Valor do rebanho bovino/pessoa ocupada;
- V18: valor da produção de leite/pessoa ocupada;
- V19: porcentagem de área ocupada segundo o produtor proprietário;
- V20: porcentagem do número de estabelecimento sem pessoal contratado;
- V21: porcentagem de empregados permanente/pessoal ocupado;
- V22: porcentagem de empregados temporários/pessoal ocupado;
- V23: porcentagem de mão de obra familiar/pessoal ocupado;
- V24: porcentagem do trabalho de empreiteira/pessoal ocupado.

**ANEXO B - Os 26 municípios retirados dos estados de Goiás e Tocantins**

- 01: Bico do Papagaio;
- 02: Araguaína;
- 03: Miracema do Norte;
- 04: Rio Formoso;
- 05: Gurupi;
- 06: Parque acional;
- 07: Jalapão;
- 08: Dianópolis;
- 09: S. Miguel do Araguaia;
- 010: Rio Vermelho;
- 011: Aragaças;
- 012: Porangatu;
- 013: C. Veadeiros;
- 014: Ceres;
- 015: Anápolis;
- 016: Iporá;
- 017: Anicuns;
- 018: Goiânia;
- 019: Paraná;
- 020: E. Brasília;
- 021: S. Goiás;
- 022: V. R. dos Bois
- 023: Meia Ponte;
- 024: Pires do Rio;
- 025: Catalão;
- 026: Quirinópolis.

**ANEXO C - Programa de Análise Fatorial para 24 variáveis retiradas dos municípios**

```
title 'Análise fatorial';
data mapeamento;
input V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17 V18 V19
      V20 V21 V22 V23 V24 @@;
cards;
151.13 28.16 5.37 4.85 33.83 15.17 2.01 0.46
0.26 59.27 0.47 4.94 289.04 21.08 4.09 921.13
10263.11 175.52 93.58 89.66 4.02 5.36 88.64 25.85
71.33 15.42 4.62 5.71 44.69 7.33 6.75 0.57
0.82 69.41 1.01 8.90 443.64 13.75 1.26 984.10
28961.89 504.07 97.40 69.23 11.51 14.17 69.06 15.43
57.52 14.61 3.94 5.49 25.95 17.66 9.93 0.33
1.51 38.59 1.02 6.18 225.54 27.18 0.46 1829.59
17296.97 479.07 92.84 65.36 10.14 15.61 67.87 16.12
19.50 8.19 2.38 4.87 21.61 9.23 26.21 0.36
4.83 25.36 1.62 4.17 255.32 76.58 7.53 10304.08
31186.97 359.33 92.47 52.14 19.22 16.71 57.72 13.68
13.66 11.08 1.23 5.26 31.26 18.39 34.70 0.39
4.73 40.24 1.29 3.60 250.17 62.50 12.63 6817.66
22567.33 289.73 91.15 57.49 15.75 14.89 67.14 14.84
37.46 14.39 2.60 5.17 17.23 8.24 14.81 0.28
2.44 17.33 0.98 4.08 159.21 31.22 0.92 2280.42
11064.69 289.51 91.38 70.17 9.35 10.81 74.05 13.28
492.48 15.04 32.74 4.48 6.02 17.66 1.08 0.23
0.19 8.29 0.81 2.50 143.22 6.94 0.15 473.81
9521.92 220.63 77.22 80.73 4.57 7.39 85.09 16.91
99.98 12.30 8.12 4.78 8.71 12.56 2.23 0.22
0.42 10.59 1.06 1.58 163.01 10.14 5.67 1289.48
```



13247.80 130.79 93.47 77.67 6.24 9.68 78.09 10.46  
15.03 9.40 1.60 5.83 52.35 5.67 30.99 0.62  
1.99 69.45 2.75 3.18 548.87 30.35 0.15 3249.73  
58208.80 335.20 96.37 47.44 20.64 21.55 53.46 17.26  
19.07 14.26 1.34 5.63 53.14 7.39 48.36 0.85  
4.76 68.95 1.73 20.03 820.39 25.81 0.55 1873.67  
57516.22 1350.04 94.69 52.85 20.76 22.65 52.60 14.30  
16.05 6.82 2.35 2.99 28.33 8.45 51.56 0.82  
5.97 65.19 1.69 17.91 368.33 29.45 1.19 4499.96  
53979.09 2130.85 54.17 47.21 17.48 24.74 55.83 14.55  
36.37 22.12 1.64 4.71 32.71 10.71 25.21 0.45  
1.48 38.84 0.85 11.27 387.09 42.67 3.93 2116.06  
17501.91 270.66 94.29 81.49 8.66 7.61 81.70 20.70  
46.35 12.76 3.63 5.04 15.62 9.82 6.96 0.28  
4.90 23.02 1.55 4.41 141.09 11.99 5.41 1377.78  
11060.35 191.22 82.89 72.09 11.02 14.23 68.34 7.46  
37.44 38.71 0.97 3.93 42.13 3.68 60.87 0.88  
9.35 38.12 0.40 51.63 861.79 142.26 108.04 6536.84  
22262.31 1058.25 93.72 74.38 8.95 14.42 71.79 21.70  
36.99 60.39 0.61 4.15 62.48 2.52 66.56 1.40  
27.99 64.08 0.30 86.39 1655.72 284.87 114.81 7226.26  
27415.30 1426.48 85.03 61.35 12.00 15.90 63.46 22.38  
33.48 21.83 1.53 5.35 63.41 2.04 46.82 1.06  
12.11 69.50 0.82 63.85 1301.49 40.89 1.70 1982.90  
59629.14 2293.42 77.26 51.26 15.51 18.16 60.42 13.75  
36.21 38.52 0.94 4.69 69.72 1.25 62.08 1.37  
8.62 77.79 0.42 87.97 1300.57 194.89 21.14 5698.20  
33763.48 1333.53 81.95 58.60 14.16 12.95 61.05 20.70  
20.98 31.66 0.66 4.06 43.13 2.92 74.64 1.17  
22.13 56.34 0.40 81.68 858.27 141.78 32.13 5947.67

```

27108.13 2059.83 76.99 44.21 22.89 18.05 52.74 16.95
68.10 23.37 2.91 6.17 30.85 6.63 6.42 0.56
1.14 45.07 1.15 3.12 501.51 13.12 2.89 697.98
21461.69 153.00 96.91 76.33 7.69 10.52 75.66 12.06
21.98 20.69 1.06 4.59 33.08 6.59 50.94 0.49
8.08 34.54 0.98 25.84 409.94 36.40 46.24 4064.82
19813.55 1118.76 92.85 51.96 16.95 19.54 58.07 15.76
7.24 9.35 0.77 7.04 54.53 3.15 52.75 0.63
28.23 57.56 2.21 16.48 607.46 53.17 123.89 19027.78
64987.42 1531.61 70.46 32.75 33.40 21.63 42.87 18.10
9.53 24.98 0.38 4.60 84.97 4.33 67.59 1.11
50.08 81.75 0.95 65.19 1029.09 125.21 164.26 11641.78
41189.08 1659.78 89.32 51.59 29.53 15.26 53.02 24.75
9.05 23.51 0.38 4.77 61.07 2.02 67.41 1.19
33.88 57.89 0.86 83.70 1175.61 127.64 152.11 12004.29
49995.36 2711.25 91.61 39.27 27.15 23.19 46.99 22.82
32.02 35.45 0.90 3.95 40.82 4.16 71.49 0.58
14.60 29.44 0.67 60.94 731.72 58.67 32.71 2621.49
20637.94 1511.36 95.28 54.34 13.98 23.26 59.66 24.16
22.15 16.06 1.38 4.24 28.20 3.31 67.26 0.57
10.43 24.12 1.14 32.17 636.29 24.63 37.23 4027.56
39622.73 1703.64 88.37 53.40 20.94 12.41 62.78 25.12
10.85 14.20 0.76 4.54 86.11 1.11 31.98 1.18
30.35 90.29 1.62 54.90 1737.53 29.09 34.42 4535.40
122357.66 3293.74 82.90 37.74 26.64 24.77 45.26 18.49
;
proc factor res data=mapeamento out=work.FactOut1
      method=prin nfact=3 rotate=varimax outstat=work.FactOut2;
var V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17 V18 V19
      V20 V21 V22 V23 V24;

```

```

title 'Análise fatorial';
proc print;
title 'Análise fatorial';
*//////////////////////////////////////;
proc iml;
RESET print ;
USE work.Factout2 ;
READ all INTO Factout2 VAR {V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14
                           V15 V16 V17 V18 V19 V20 V21 V22 V23 V24};
Xb = t(Factout2[1,]);
USE work.Factout1 ;
READ all INTO Factout1 VAR {V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14
                           V15 V16 V17 V18 V19 V20 V21 V22 V23 V24};
X = Factout1[,1:24];
XT = t(X);
NL = nrow(X);
NC = ncol(X);
UM = j(NL,1,1);
UMT = t(UM);
z1 = ((XT[1,]) - Xb[1,1]*UMT)/(Factout2[2,1]);
z2 = ((XT[2,]) - Xb[2,1]*UMT)/(Factout2[2,2]);
z3 = ((XT[3,]) - Xb[3,1]*UMT)/(Factout2[2,3]);
z4 = ((XT[4,]) - Xb[4,1]*UMT)/(Factout2[2,4]);
z5 = ((XT[5,]) - Xb[5,1]*UMT)/(Factout2[2,5]);
z6 = ((XT[6,]) - Xb[6,1]*UMT)/(Factout2[2,6]);
z7 = ((XT[7,]) - Xb[7,1]*UMT)/(Factout2[2,7]);
z8 = ((XT[8,]) - Xb[8,1]*UMT)/(Factout2[2,8]);
z9 = ((XT[9,]) - Xb[9,1]*UMT)/(Factout2[2,9]);
z10 = ((XT[10,]) - Xb[10,1]*UMT)/(Factout2[2,10]);
z11 = ((XT[11,]) - Xb[11,1]*UMT)/(Factout2[2,11]);

```

```

z12 = ((XT[12,]) - Xb[12,1]*UMT)/(Factout2[2,12]);
z13 = ((XT[13,]) - Xb[13,1]*UMT)/(Factout2[2,13]);
z14 = ((XT[14,]) - Xb[14,1]*UMT)/(Factout2[2,14]);
z15 = ((XT[15,]) - Xb[15,1]*UMT)/(Factout2[2,15]);
z16 = ((XT[16,]) - Xb[16,1]*UMT)/(Factout2[2,16]);
z17 = ((XT[17,]) - Xb[17,1]*UMT)/(Factout2[2,17]);
z18 = ((XT[18,]) - Xb[18,1]*UMT)/(Factout2[2,18]);
z19 = ((XT[19,]) - Xb[19,1]*UMT)/(Factout2[2,19]);
z20 = ((XT[20,]) - Xb[20,1]*UMT)/(Factout2[2,20]);
z21 = ((XT[21,]) - Xb[21,1]*UMT)/(Factout2[2,21]);
z22 = ((XT[22,]) - Xb[22,1]*UMT)/(Factout2[2,22]);
z23 = ((XT[23,]) - Xb[23,1]*UMT)/(Factout2[2,23]);
z24 = ((XT[24,]) - Xb[24,1]*UMT)/(Factout2[2,24]);

Zj = t(z1//z2//z3//z4//z5//z6//z7//z8//z9//z10//z11//z12//z13//z14//
      z15//z16//z17//z18//z19//z20//z21//z22//z23//z24);
create arq from Zj;
append from Zj;
quit;
run;
title 'Análise fatorial padr';
data MP (rename=(col1=z1 col2=z2 col3=z3 col4=z4 col5=z5 col6=z6 col7=z7
col8=z8 col9=z9 col10=z10 col11=z11 col12=z12 col13=z13 col14=z14
col15=z15 col16=z16 col17=z17 col18=z18 col19=z19 col20=z20 col21=z21
col22=z22 col23=z23 col24=z24));
set arq;
run;
*//////////////////////////////////////;
proc factor res data=MP out=work.FactOut1_
      method=prin nfact=3 rotate=varimax scree plot

```

```

        outstat=work.FactOut2_;
var z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14 z15 z16 z17 z18 z19
z20 z21 z22 z23 z24;
title 'Análise fatorial padr';
proc print;
title 'Análise fatorial padr';

proc iml;
RESET print;

USE work.FactOut2_;
READ all INTO FactOut2_ VAR {z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14
                            z15 z16 z17 z18 z19 z20 z21 z22 z23 z24};

Lz = t(FactOut2_[61:63,]);
LzT = FactOut2_[61:63,];
USE work.FactOut1_ ;
READ all INTO FactOut1_ VAR {z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14
                            z15 z16 z17 z18 z19 z20 z21 z22 z23 z24};

Zj = t(FactOut1_[,1:24]);
f_z1 = t(inv(LzT*Lz)*(LzT*Zj[,1]));
f_z2 = t(inv(LzT*Lz)*(LzT*Zj[,2]));
f_z3 = t(inv(LzT*Lz)*(LzT*Zj[,3]));
f_z4 = t(inv(LzT*Lz)*(LzT*Zj[,4]));
f_z5 = t(inv(LzT*Lz)*(LzT*Zj[,5]));
f_z6 = t(inv(LzT*Lz)*(LzT*Zj[,6]));
f_z7 = t(inv(LzT*Lz)*(LzT*Zj[,7]));
f_z8 = t(inv(LzT*Lz)*(LzT*Zj[,8]));
f_z9 = t(inv(LzT*Lz)*(LzT*Zj[,9]));
f_z10 = t(inv(LzT*Lz)*(LzT*Zj[,10]));
f_z11 = t(inv(LzT*Lz)*(LzT*Zj[,11]));

```

```
f_z12 = t(inv(LzT*Lz)*(LzT*Zj[,12]));
f_z13 = t(inv(LzT*Lz)*(LzT*Zj[,13]));
f_z14 = t(inv(LzT*Lz)*(LzT*Zj[,14]));
f_z15 = t(inv(LzT*Lz)*(LzT*Zj[,15]));
f_z16 = t(inv(LzT*Lz)*(LzT*Zj[,16]));
f_z17 = t(inv(LzT*Lz)*(LzT*Zj[,17]));
f_z18 = t(inv(LzT*Lz)*(LzT*Zj[,18]));
f_z19 = t(inv(LzT*Lz)*(LzT*Zj[,19]));
f_z20 = t(inv(LzT*Lz)*(LzT*Zj[,20]));
f_z21 = t(inv(LzT*Lz)*(LzT*Zj[,21]));
f_z22 = t(inv(LzT*Lz)*(LzT*Zj[,22]));
f_z23 = t(inv(LzT*Lz)*(LzT*Zj[,23]));
f_z24 = t(inv(LzT*Lz)*(LzT*Zj[,24]));
f_z25 = t(inv(LzT*Lz)*(LzT*Zj[,25]));
f_z26 = t(inv(LzT*Lz)*(LzT*Zj[,26]));

EF = f_z1//f_z2//f_z3//f_z4//f_z5//f_z6//f_z7//f_z8//f_z9//f_z10//f_z11//
      f_z12//f_z13//f_z14//f_z15//f_z16//f_z17//f_z18//f_z19//f_z20//
      f_z21//f_z22//f_z23//f_z24//f_z25//f_z26;

create ESCORE from EF;

append from EF;

quit;

run;

proc corr data=ESCORE;

run;
```

## ANEXO D - Programa para a obtenção do gráfico Biplot

```

proc iml;
RESET print;
USE work.FactOut2_;
READ all INTO FactOut2_ VAR {z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14
                             z15 z16 z17 z18 z19 z20 z21 z22 z23 z24};

Lz = t(FactOut2_[61:63,]);
USE work.FactOut1_ ;
READ all INTO FactOut1_ VAR {z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13 z14
                             z15 z16 z17 z18 z19 z20 z21 z22 z23 z24};

Xz = t(FactOut1_[,1:24]);
XL = Xz||Lz;
XLXLL = XL*t(XL);
XLLXL = t(XL)*XL;
call svd(UXL,BXL,VXL,XL);
DBXL = diag(sqrt(BXL));
*Biplot SQRT;
GXL = UXL*DBXL;
HXL = DBXL*t(VXL);
YXL = GXL//t(HXL);
run;

data BYXL_var_fat;
input map_VAR_FAT $ X Y Z;
cards;
V1 -0.699392 0.4207336 1.1518473 0.1804395 0.8931823 0.048614 0.4439346
0.3473878 -0.10814 -0.058536 0.085528 -0.097842 -0.093591 -0.067375
-0.020226 0.1458326 0.0958084 0.0478809 -0.155057 -0.052375 -0.042764
0.2177028 -0.143963 0.006422 0 0 0 0 0
V2 0.4476048 1.3460715 -0.287841 -0.150404 -0.040131 -0.314744 0.2019436

```

0.3587519 -0.128206 0.182496 0.4371479 0.2952499 0.4458496 -0.189272  
0.0084452 0.3880819 0.1046158 -0.019386 -0.231633 0.0958866 0.0232004  
-0.032957 0.0977418 0.0247869 0 0 0 0 0  
V3 -0.711973 0.2172059 1.2434705 0.2556124 0.7920422 -0.011627 0.5455463  
0.337235 0.013141 0.0326396 -0.066865 -0.211585 -0.117625 0.0436292  
-0.06172 -0.024777 -0.09668 -0.028065 0.1658755 0.0396579 0.0537213  
-0.215068 0.1378682 0.0114876 0 0 0 0 0  
V4 -0.186052 -0.775958 -0.929331 0.1089322 1.1025844 -0.665047 0.4289792  
-0.024523 -0.508047 -0.665049 0.2863049 0.3614407 0.1690991 0.1489538  
0.1571197 -0.081068 0.0177998 -0.017694 0.1366144 0.040642 -0.002311  
0.0041474 -0.017079 0.0060246 0 0 0 0 0  
V5 1.0660265 0.0652058 -0.268395 -0.475665 0.7449811 0.0860439 -0.179064  
0.0472716 0.1680561 -0.074137 0.0397109 -0.209348 0.0649646 0.055953  
-0.605977 -0.219387 -0.157308 -0.107916 -0.032062 0.0037378 -0.038573  
0.1914103 0.112967 0.0164416 0 0 0 0 0  
V6 -1.005239 -0.095101 0.1874667 0.3129 0.1876452 0.1183767 -0.654654  
0.4904467 0.9928918 -0.516255 -0.185001 0.4269351 0.293933 0.3766461  
0.0597116 -0.045743 0.0855196 -0.037752 -0.102361 0.0099687 -0.021137  
-0.016461 0.0151631 0.0006951 0 0 0 0 0  
V7 1.0243476 0.3247879 0.078844 0.3667499 -0.678436 0.239609 0.0949628  
0.4021281 -0.465588 -0.313389 0.0814735 -0.257743 0.1018615 0.747377  
-0.291562 0.192785 0.0186812 0.0711846 0.0909895 0.0119972 -0.023878  
-0.049329 -0.06332 0.031734 0 0 0 0 0  
V8 1.0779238 0.4772369 0.1464109 -0.509885 0.1738519 -0.382551 -0.106941  
0.1165642 0.0349056 -0.022303 -0.124298 -0.215083 -0.282947 0.2483272  
0.3858189 0.0213303 0.2385625 0.1383205 0.0534549 0.0110556 -0.24132  
0.060679 0.0994197 -0.060417 0 0 0 0 0  
V9 1.0312859 0.1415502 0.1294096 0.6163626 0.442731 0.109107 0.0321959  
-0.601105 0.5899243 0.066432 0.1228721 -0.328487 0.6429671 -0.233586  
-0.019965 0.077393 0.0884682 0.1773028 0.2190961 0.0056805 -0.093797



-0.035353 -0.052591 0.0168987 0 0 0 0 0  
V10 0.8684325 -0.14777 -0.158979 -0.86141 0.7141956 -0.144509 -0.849254  
0.3042383 0.1947103 -0.210328 0.3370625 -0.393122 -0.333168 -0.124264  
0.0738575 0.1821738 -0.056053 0.0965492 -0.044715 -0.028285 0.1119296  
-0.14483 -0.07775 0.0347623 0 0 0 0 0  
V11 0.0079337 -1.431614 -0.084828 0.0649861 0.1933061 0.0090712 -0.093538  
0.5518104 -0.241997 0.8394414 0.0328059 -0.243393 0.3533681 0.1739233  
0.1711412 -0.266876 0.1839956 0.113626 -0.10868 -0.059426 0.0275747  
0.012697 0.0292742 0.0547174 0 0 0 0 0  
V12 1.0021323 0.7774332 0.1949901 -0.159641 -0.157175 -0.12686 0.3814207  
-0.073001 0.0444764 -0.375346 -0.139409 -0.111246 0.0039797 -0.088192  
-0.039255 -0.587413 0.2905738 0.0197029 -0.178183 0.1525206 0.0989661  
-0.089777 -0.057525 0.002048 0 0 0 0 0  
V13 1.0428226 0.4260244 0.0711864 -0.632272 0.3979571 -0.108172 0.3445499  
-0.080646 0.0389892 0.2597193 -0.30466 0.4014072 0.2548945 0.2123449  
-0.057435 -0.019986 -0.076946 -0.007387 -0.048361 -0.34225 0.0626198  
-0.102983 -0.044293 -0.062245 0 0 0 0 0  
V14 0.7414091 0.9535043 -0.20287 0.3105289 -0.120035 -0.820632 -0.097914  
0.6544148 0.1873421 0.2975846 -0.511673 0.0479515 0.0032098 -0.08001  
0.1899794 -0.065886 -0.179379 -0.093066 0.249486 0.0521032 0.1006828  
0.1134875 -0.050849 0.038356 0 0 0 0 0  
V15 0.8372776 0.3790801 -0.175421 1.1141272 0.4148936 0.0486827 0.064985  
-0.358457 0.318161 0.4735416 0.5638055 0.0407923 -0.317681 0.4025421  
0.2101151 -0.119318 -0.242384 -0.076282 -0.158735 0.0658248 0.0278978  
-0.005154 -0.015364 -0.023106 0 0 0 0 0  
V16 0.7832094 -0.277247 -0.129848 1.3057391 0.2438054 -0.423211 -0.33832  
0.0123647 -0.144971 0.0555897 -0.309387 0.4069187 -0.445808 -0.22499  
-0.351062 0.0943174 0.2508684 0.1318726 -0.05895 -0.095984 -0.031238  
-0.056909 0.0289537 0.0206587 0 0 0 0 0  
V17 0.8345824 -0.78916 0.3579588 -0.631811 0.4853019 0.3114187 -0.034059

-0.165908 -0.050544 0.4665952 -0.416049 0.511161 -0.010236 0.1285525  
-0.102939 0.1977041 0.0229847 -0.058755 0.0188572 0.321718 -0.034924  
-0.039808 -0.065278 -0.008565 0 0 0 0 0

V18 1.058087 -0.088794 0.6762939 -0.316284 -0.168099 0.3821853 0.146088  
-0.544922 0.0714711 -0.264372 0.0290833 0.4111232 -0.154355 0.1062627  
0.2947773 0.0301843 0.0446408 0.0028113 0.0384515 -0.097395 0.0827431  
0.095325 0.068394 0.1240554 0 0 0 0 0

V19 -0.341256 0.2804231 -1.434768 -0.0766 0.3270511 0.8674997 0.8879664  
0.2748326 0.4939356 0.0597271 -0.352892 -0.146239 -0.293576 0.0133632  
0.0309598 0.128082 0.0728345 0.1079442 -0.01112 -0.005642 0.0012643  
-0.013344 0.0009907 0.0312401 0 0 0 0 0

V20 -0.981433 0.7781117 -0.250084 -0.222584 0.2837853 -0.017629 -0.281048  
-0.314368 0.0293634 0.4063728 0.2473782 -0.067855 -0.122448 0.2090321  
-0.11488 0.0068868 0.4519562 -0.356374 0.1887018 -0.054951 0.0329394  
-0.010827 -0.027988 0.0169371 0 0 0 0 0

V21 0.97712 -0.757942 0.0391341 0.5337256 0.1401274 0.2070184 0.0077079  
-0.05677 -0.043702 -0.292161 -0.306644 -0.433631 0.1308156 -0.004003  
0.1512732 0.2901364 0.1666042 -0.190037 -0.049876 0.0123271 0.246328  
0.0882797 0.056155 -0.070541 0 0 0 0 0

V22 0.8493778 -0.780855 0.2452731 -0.16655 -0.532282 0.162647 0.4142131  
0.6396566 0.4778204 0.0610305 0.7312 0.3793524 -0.176818 -0.174284  
-0.131444 -0.052819 0.1357763 0.0402826 0.2098541 -0.007568 0.0632282  
0.0507519 -0.005966 -0.052507 0 0 0 0 0

V23 -1.027982 0.7498118 -0.062537 -0.153889 0.2235948 0.0614131 -0.339484  
-0.308067 -0.143224 0.1694686 0.0244536 0.1589063 0.0061552 0.2212255  
-0.096037 -0.030154 0.0459194 0.4497406 0.1012778 0.0707292 0.2273902  
0.0993623 0.0386513 -0.035884 0 0 0 0 0

V24 0.5498488 0.7993459 -0.077832 0.3927548 0.4301952 1.3206754 -0.584688  
0.4229898 -0.625888 -0.088167 0.039792 0.25013 0.1021727 -0.246778  
0.192043 -0.1903 -0.022538 -0.053014 0.101174 -0.020857 -0.041857 -0.01203

```

0.0022159 -0.022132 0 0 0 0 0
F1 0.6458627 -0.369546 0.1041262 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2.3648E-9 0 0 0 0
F2 0.4843539 0.5154379 -0.018174 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-3.92E-10 0 0 0 0
F3 -0.115805 0.0948001 0.5047157 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.068E-10 0 0 0 0
;
data ESCORE(rename=(col1=f1 col2=f2 col3=f3));
set ESCORE;
run;
proc iml;
RESET print;
USE work.ESCORE;
READ all INTO ESCORE VAR {f1 f2 f3};
E = ESCORE[,1:3];
EEL = E*t(E);
ELE = t(E)*E;
call svd(UE,BE,VE,E);
DBE = diag(sqrt(BE));
G = UE*DBE;
H = DBE*t(VE);
YE = G//t(H);
run;
data BYE_obj_fat;
input map_MUNI_FAT $ X Y Z;
cards;
01 0.1304576 -0.025033 -0.774079
02 -0.208026 -0.334467 -0.220533
03 -0.331319 -0.182644 -0.287133

```

```
04 -0.410465 -0.254957 0.065675
05 -0.324582 -0.301835 -0.120419
06 -0.271465 -0.226102 -0.394668
07 -0.409547 1.6472521 -0.57641
08 -0.385445 0.049555 -0.503713
09 -0.520973 -0.551583 0.3562226
010 -0.283572 -0.301619 0.3311994
011 -0.412641 0.9287503 0.5970841
012 0.0440674 -0.262943 -0.546612
013 -0.498468 -0.04253 -0.208446
014 0.6435789 -0.08301 -0.361484
015 1.1508647 0.0513247 -0.16793
016 0.0987287 0.2126174 0.3181192
017 0.7210765 0.0429336 -0.039833
018 0.4483533 0.3485964 0.241118
019 -0.196866 -0.472023 -0.416693
020 -0.071026 -0.139355 0.0085948
021 -0.391294 -0.262658 0.8816202
022 0.5924323 -0.178882 0.3679279
023 0.493241 -0.033889 0.5199815
024 0.3608515 0.0209694 -0.070006
025 0.049887 0.0819659 0.0574789
26 -0.01785 0.269565 0.9429352
F1 0 0 2.236068
F2 2.2294843 0.1714632 0
F3 -0.171463 2.2294843 0
;
data BYXL_var_obj_fat;
input map_VAR_MUNI_FAT $ X Y Z;
cards;
```

V1 -0.699392 0.4207336 1.1518473 0.1804395 0.8931823 0.048614 0.4439346  
0.3473878 -0.10814 -0.058536 0.085528 -0.097842 -0.093591 -0.067375  
-0.020226 0.1458326 0.0958084 0.0478809 -0.155057 -0.052375 -0.042764  
0.2177028 -0.143963 0.006422 0 0 0 0 0

V2 0.4476048 1.3460715 -0.287841 -0.150404 -0.040131 -0.314744 0.2019436  
0.3587519 -0.128206 0.182496 0.4371479 0.2952499 0.4458496 -0.189272  
0.0084452 0.3880819 0.1046158 -0.019386 -0.231633 0.0958866 0.0232004  
-0.032957 0.0977418 0.0247869 0 0 0 0 0

V3 -0.711973 0.2172059 1.2434705 0.2556124 0.7920422 -0.011627 0.5455463  
0.337235 0.013141 0.0326396 -0.066865 -0.211585 -0.117625 0.0436292  
-0.06172 -0.024777 0.09668 -0.028065 0.1658755 0.0396579 0.0537213  
-0.215068 0.1378682 0.0114876 0 0 0 0 0

V4 -0.186052 -0.775958 -0.929331 0.1089322 1.1025844 -0.665047 0.4289792  
-0.024523 -0.508047 -0.665049 0.2863049 0.3614407 0.1690991 0.1489538  
0.1571197 -0.081068 0.0177998 -0.017694 0.1366144 0.040642 -0.002311  
0.0041474 -0.017079 0.0060246 0 0 0 0 0

V5 1.0660265 0.0652058 -0.268395 -0.475665 0.7449811 0.0860439 -0.179064  
0.0472716 0.1680561 -0.074137 0.0397109 -0.209348 0.0649646 0.055953  
-0.605977 -0.219387 -0.157308 -0.107916 -0.032062 0.0037378 -0.038573  
0.1914103 0.112967 0.0164416 0 0 0 0 0

V6 -1.005239 -0.095101 0.1874667 0.3129 0.1876452 0.1183767 -0.654654  
0.4904467 0.9928918 -0.516255 -0.185001 0.4269351 0.293933 0.3766461  
0.0597116 -0.045743 0.0855196 -0.037752 -0.102361 0.0099687 -0.021137  
-0.016461 0.0151631 0.0006951 0 0 0 0 0

V7 1.0243476 0.3247879 0.078844 0.3667499 -0.678436 0.239609 0.0949628  
0.4021281 -0.465588 -0.313389 0.0814735 -0.257743 0.1018615 0.747377  
-0.291562 0.192785 0.0186812 0.0711846 0.0909895 0.0119972 -0.023878  
-0.049329 -0.06332 0.031734 0 0 0 0 0

V8 1.0779238 0.4772369 0.1464109 -0.509885 0.1738519 -0.382551 -0.106941  
0.1165642 0.0349056 -0.022303 -0.124298 -0.215083 -0.282947 0.2483272

0.3858189 0.0213303 0.2385625 0.1383205 0.0534549 0.0110556 -0.24132  
0.060679 0.0994197 -0.060417 0 0 0 0 0  
V9 1.0312859 0.1415502 0.1294096 0.6163626 0.442731 0.109107 0.0321959  
-0.601105 0.5899243 0.066432 0.1228721 -0.328487 0.6429671 -0.233586  
-0.019965 0.077393 0.0884682 0.1773028 0.2190961 0.0056805 -0.093797  
-0.035353 -0.052591 0.0168987 0 0 0 0 0  
V10 0.8684325 -0.14777 -0.158979 -0.86141 0.7141956 -0.144509 -0.849254  
0.3042383 0.1947103 -0.210328 0.3370625 -0.393122 -0.333168 -0.124264  
0.0738575 0.1821738 -0.056053 0.0965492 -0.044715 -0.028285 0.1119296  
-0.14483 -0.07775 0.0347623 0 0 0 0 0  
V11 0.0079337 -1.431614 -0.084828 0.0649861 0.1933061 0.0090712 -0.093538  
0.5518104 -0.241997 0.8394414 0.0328059 -0.243393 0.3533681 0.1739233  
0.1711412 -0.266876 0.1839956 0.113626 -0.10868 -0.059426 0.0275747  
0.012697 0.0292742 0.0547174 0 0 0 0 0  
V12 1.0021323 0.7774332 0.1949901 -0.159641 -0.157175 -0.12686 0.3814207  
-0.073001 0.0444764 -0.375346 -0.139409 -0.111246 0.0039797 -0.088192  
-0.039255 -0.587413 0.2905738 0.0197029 -0.178183 0.1525206 0.0989661  
-0.089777 -0.057525 0.002048 0 0 0 0 0  
V13 1.0428226 0.4260244 0.0711864 -0.632272 0.3979571 -0.108172 0.3445499  
-0.080646 0.0389892 0.2597193 -0.30466 0.4014072 0.2548945 0.2123449  
-0.057435 -0.019986 -0.076946 -0.007387 -0.048361 -0.34225 0.0626198  
-0.102983 -0.044293 -0.062245 0 0 0 0 0  
V14 0.7414091 0.9535043 -0.20287 0.3105289 -0.120035 -0.820632 -0.097914  
0.6544148 0.1873421 0.2975846 -0.511673 0.0479515 0.0032098 -0.08001  
0.1899794 -0.065886 -0.179379 -0.093066 0.249486 0.0521032 0.1006828  
0.1134875 -0.050849 0.038356 0 0 0 0 0  
V15 0.8372776 0.3790801 -0.175421 1.1141272 0.4148936 0.0486827 0.064985  
-0.358457 0.318161 0.4735416 0.5638055 0.0407923 -0.317681 0.4025421  
0.2101151 -0.119318 -0.242384 -0.076282 -0.158735 0.0658248 0.0278978  
-0.005154 -0.015364 -0.023106 0 0 0 0 0

V16 0.7832094 -0.277247 -0.129848 1.3057391 0.2438054 -0.423211 -0.33832  
0.0123647 -0.144971 0.0555897 -0.309387 0.4069187 -0.445808 -0.22499  
-0.351062 0.0943174 0.2508684 0.1318726 -0.05895 -0.095984 -0.031238  
-0.056909 0.0289537 0.0206587 0 0 0 0 0

V17 0.8345824 -0.78916 0.3579588 -0.631811 0.4853019 0.3114187 -0.034059  
-0.165908 -0.050544 0.4665952 -0.416049 0.511161 -0.010236 0.1285525  
-0.102939 0.1977041 0.0229847 -0.058755 0.0188572 0.321718 -0.034924  
-0.039808 -0.065278 -0.008565 0 0 0 0 0

V18 1.058087 -0.088794 0.6762939 -0.316284 -0.168099 0.3821853 0.146088  
-0.544922 0.0714711 -0.264372 0.0290833 0.4111232 -0.154355 0.1062627  
0.2947773 0.0301843 0.0446408 0.0028113 0.0384515 -0.097395 0.0827431  
0.095325 0.068394 0.1240554 0 0 0 0 0

V19 -0.341256 0.2804231 -1.434768 -0.0766 0.3270511 0.8674997 0.8879664  
0.2748326 0.4939356 0.0597271 -0.352892 -0.146239 -0.293576 0.0133632  
0.0309598 0.128082 0.0728345 0.1079442 -0.01112 -0.005642 0.0012643  
-0.013344 0.0009907 0.0312401 0 0 0 0 0

V20 -0.981433 0.7781117 -0.250084 -0.222584 0.2837853 -0.017629 -0.281048  
-0.314368 0.0293634 0.4063728 0.2473782 -0.067855 -0.122448 0.2090321  
-0.11488 0.0068868 0.4519562 -0.356374 0.1887018 -0.054951 0.0329394  
-0.010827 -0.027988 0.0169371 0 0 0 0 0

V21 0.97712 -0.757942 0.0391341 0.5337256 0.1401274 0.2070184 0.0077079  
-0.05677 -0.043702 -0.292161 -0.306644 -0.433631 0.1308156 -0.004003  
0.1512732 0.2901364 0.1666042 -0.190037 -0.049876 0.0123271 0.246328  
0.0882797 0.056155 -0.070541 0 0 0 0 0

V22 0.8493778 -0.780855 0.2452731 -0.16655 -0.532282 0.162647 0.4142131  
0.6396566 0.4778204 0.0610305 0.7312 0.3793524 -0.176818 -0.174284  
-0.131444 -0.052819 0.1357763 0.0402826 0.2098541 -0.007568 0.0632282  
0.0507519 -0.005966 -0.052507 0 0 0 0 0

V23 -1.027982 0.7498118 -0.062537 -0.153889 0.2235948 0.0614131 -0.339484  
-0.308067 -0.143224 0.1694686 0.0244536 0.1589063 0.0061552 0.2212255

-0.096037 -0.030154 0.0459194 0.4497406 0.1012778 0.0707292 0.2273902  
0.0993623 0.0386513 -0.035884 0 0 0 0 0  
V24 0.5498488 0.7993459 -0.077832 0.3927548 0.4301952 1.3206754 -0.584688  
0.4229898 -0.625888 -0.088167 0.039792 0.25013 0.1021727 -0.246778 0.192043  
-0.1903 -0.022538 -0.053014 0.101174 -0.020857 -0.041857 -0.01203 0.0022159  
-0.022132 0 0 0 0 0  
01 -0.970069 0.7799376 -0.224898 -0.224631 0.6182168 0.5322924 -0.902362  
-0.023954 -0.129501 -0.182199 0.2018287 0.2124588 -0.035949 -0.295059  
0.1519753 0.1006947 0.1832548 0.0404654 -0.20182 -0.057084 -0.002748  
0.0842448 -0.083309 0.0563811 1.9734E-9 0 0 0 0  
02 -0.488683 -0.153692 -0.400235 -0.529248 0.4545668 -0.002635 0.0354267  
-0.037959 0.0447122 -0.201175 0.2478029 -0.258447 -0.502012 -0.31204  
0.0107509 0.1156407 -0.270555 -0.055302 0.1722951 -0.012718 0.196736  
-0.15984 -0.044579 -0.041668 1.8718E-9 0 0 0 0  
03 -0.747045 -0.192831 -0.229709 0.005272 0.0038611 0.0605881 -0.146795  
0.2219013 0.5318994 -0.508463 0.0950639 0.4370133 0.160997 -0.046385  
0.252351 -0.241012 -0.145854 -0.261427 0.057794 0.0728834 -0.149237  
-0.074554 0.0169965 -0.006994 -1.75E-10 0 0 0 0  
04 -0.311896 -0.595707 -0.218653 0.5666614 -0.366999 -0.176791 0.0945273  
0.2590169 0.0888832 0.2326422 -0.732647 0.0659177 -0.27406 -0.448793  
-0.186677 0.1700328 0.1215694 0.0272435 0.023853 -0.211349 -0.049963  
-0.022046 -0.039432 0.0133115 1.463E-10 0 0 0 0  
05 -0.477994 -0.358685 -0.325497 0.3746806 -0.149037 -0.114561 -0.43903  
0.3645751 0.5222445 -0.418943 -0.387522 0.2623269 0.1567615 0.4140485  
-0.203433 0.0382838 -0.072161 0.2204452 -0.066824 0.0356803 0.0805933  
0.029881 -0.049796 -0.070368 2.5251E-9 0 0 0 0  
06 -0.828911 -0.049189 -0.311117 0.1113902 -0.334108 -0.166275 0.1529743  
-0.454167 -0.139722 0.0429124 -0.113342 -0.023378 0.0651311 -0.173437  
0.0221121 -0.083777 -0.225867 0.050328 0.3117265 0.0926086 -0.076133  
0.2736949 -0.035866 0.0347612 2.2093E-9 0 0 0 0



07 -1.571326 0.5060636 1.7662761 0.4508218 0.9312483 0.0237927 0.4916021  
0.3527881 -0.066963 -0.011618 -0.050742 -0.107179 0.0006728 0.0736565  
-0.062272 0.0217225 0.021226 0.0032931 0.0290895 -0.002955 0.008477  
0.0440852 0.0152505 -0.017723 1.2689E-9 0 0 0 0  
08 -1.160018 -0.005587 -0.012699 0.1061633 -0.18579 -0.153457 0.2871778  
-0.451367 0.3023806 0.2980289 -0.227794 0.011063 -0.034713 0.1155931  
0.0071018 -0.115672 0.0577313 0.068276 -0.074642 0.1208036 -0.033897  
-0.248069 0.0217941 0.0973941 2.6767E-9 0 0 0 0  
09 0.0381344 -1.049899 -0.469818 -0.35581 0.313375 0.2036393 0.0895891  
0.9187835 -0.281504 0.4949697 0.0193325 -0.403996 0.1283205 -0.024943  
0.1515151 -0.282699 -0.058648 0.1832022 -0.097757 0.1399199 0.0342813  
0.0304032 -0.017854 0.0217541 6.431E-10 0 0 0 0  
010 0.2192915 -0.685346 -0.225314 -0.598155 0.0991518 0.105999 0.3069666  
0.4747851 0.1255012 -0.178845 0.1005423 -0.117605 -0.070426 0.5172466  
0.0576058 0.3104125 0.2577845 -0.229058 0.1772054 -0.101552 -0.030523  
0.0500623 -0.011973 0.0565628 .9382E-9 0 0 0 0  
011 0.2458592 -0.682494 1.2336384 -0.288247 -1.005974 -0.120862 -1.020659  
0.0856021 0.0460769 0.2960783 0.3796797 -0.051642 -0.218726 0.0264233  
0.0782663 -0.008278 -0.041715 0.0955114 0.0938843 -0.049057 -0.072918  
-0.035423 0.0315714 -0.005612 1.4083E-9 0 0 0 0  
012 -0.694901 0.4227935 -0.427927 -0.106974 -0.003773 0.3100445 -0.496259  
-0.206673 -0.231807 0.11334 -0.210785 0.0098983 0.1017112 0.1395622  
-0.148777 -0.034025 0.1615212 -0.013197 0.1474058 0.0209535 0.1887311  
0.0203726 0.2100691 -0.001162 5.194E-10 0 0 0 0  
013 -0.84176 -0.40522 -0.032922 0.0324532 -0.408989 -0.575391 0.1356281  
-0.410154 0.2338107 0.3277299 0.2747581 -0.298124 0.3226467 -0.058524  
0.0460063 -0.235588 0.2091544 -0.260557 -0.157694 -0.105303 0.1346174  
0.0682283 -0.059019 -0.05204 1.5397E-9 0 0 0 0  
014 0.2002397 1.0029383 -0.259087 0.2369676 -0.272892 0.0888161 0.0472006  
-0.03457 -0.086661 0.5846233 0.1932548 0.2437453 -0.543526 0.3688426

-0.223469 0.0505203 0.0339367 -0.109399 -0.042069 0.1917313 -0.020592  
0.0323101 -0.099821 -0.024841 4.724E-10 0 0 0 0  
015 1.0141113 1.4579711 -0.127734 0.0164726 0.0681329 -0.702239 0.026583  
0.4246642 0.1930593 0.5410164 0.0524488 0.3144784 0.4337487 -0.031459  
0.2425861 0.1678713 -0.192816 0.0315142 0.0770836 -0.111736 0.0758718  
-0.037133 0.0046837 0.0224452 1.5918E-9 0 0 0 0  
016 0.5307621 -0.087292 0.3001108 -0.916673 0.0822421 -0.379707 0.2309791  
-0.477963 -0.243982 -0.455031 0.0734734 0.1873053 0.0724087 0.253893  
-0.25888 -0.211108 -0.077965 0.2089015 -0.031006 -0.133433 0.0615569  
-0.007083 -0.073088 0.0497503 -1.32E-10 0 0 0 0  
017 0.7279336 0.8626955 -0.052716 -0.553379 0.0808205 -0.619388 -0.193807  
0.3462783 -0.477998 -0.278395 -0.422642 -0.193399 -0.219699 -0.085545  
-0.187482 -0.332318 0.0032144 -0.194072 -0.091441 -0.008747 -0.128182  
-0.000414 0.0540607 -0.020028 2.602E-9 0 0 0 0  
018 0.7802303 0.4131852 0.3903467 0.0103596 -0.630265 -0.382631 0.0967189  
-0.04231 -0.009599 -0.623975 -0.195817 -0.409169 0.1054475 -0.209302  
0.24629 0.3063133 0.1963825 0.0493501 -0.036183 0.2594726 0.0733357  
0.0000102 -0.023508 0.0051872 6.491E-10 0 0 0 0  
019 -0.735613 -0.020186 -0.601682 -0.419452 0.3449812 -0.40927 0.3005612  
-0.331002 -0.218762 0.096264 0.2696813 -0.038091 0.0226453 0.0521381  
0.1637249 0.2556847 0.123185 0.2518628 -0.037418 -0.037607 -0.229256  
-0.033955 0.0640492 -0.084256 .2583E-9 0 0 0 0  
020 -0.04129 -0.129338 -0.14535 0.2108794 -0.549945 0.2157295 0.3808922  
0.0401473 0.1359721 -0.175203 0.2879561 -0.112639 -0.147872 -0.011509  
-0.122849 0.2434336 -0.361524 -0.059387 -0.344385 -0.056581 0.0082819  
0.1048027 0.1225962 0.0348873 1.3842E-9 0 0 0 0  
021 0.8889179 -1.250537 -0.039751 1.1964687 0.6364817 -0.559729 -0.180386  
-0.124203 -0.689387 -0.11131 0.3315107 0.4336303 0.0840775 -0.069109  
-0.107438 0.0792658 0.0322192 -0.09068 0.0013209 0.0782401 0.0289983  
-0.041048 0.0171324 0.0163154 1.5218E-9 0 0 0 0

90

022 1.2141947 0.313523 -0.190339 0.7453841 0.6998952 0.3123928 -0.424722  
-0.353758 .5111363 0.0131943 0.0945421 -0.825267 0.1973889 0.0876112  
-0.230945 -0.024426 -0.061033 0.0051866 0.064118 -0.043898 -0.11424  
-0.031857 -0.01901 0.0102174 9.192E-10 0 0 0 0

023 1.2993137 0.1189615 0.0203697 0.6063005 0.1408075 0.2847574 0.4234821  
-0.162084 0.3541192 -0.142124 0.031456 0.1791981 -0.512073 0.0228991  
0.4431753 -0.349672 0.1670287 0.1522939 0.0123578 -0.088056 0.0599362  
0.071659 0.0502899 -0.010579 1.658E-9 0 0 0 0

024 0.2920599 0.4729438 -0.038619 0.0525766 -0.700632 0.809493 0.5510248  
0.3597649 -0.086422 -0.162613 0.4957613 0.1837141 0.3269271 -0.343267  
-0.388937 -0.111322 0.1764015 0.0761709 0.10653 -0.004042 -0.028539  
-0.073181 -0.008656 -0.019219 2.034E-9 0 0 0 0

025 0.123517 0.0335887 0.0990216 0.2843538 -0.438989 0.910055 0.0474288  
-0.302735 -0.828359 0.0027118 -0.42311 -0.077899 0.2281039 0.2704653  
0.3407066 0.0363721 -0.178488 -0.097565 -0.026146 -0.098047 -0.007643  
-0.092495 -0.060381 -0.033163 1.5071E-9 0 0 0 0

026 1.2949398 -0.7186 0.5243043 -1.014635 0.5736125 0.5053364 0.1052562  
-0.435409 0.4008712 0.4063818 -0.384691 0.376085 0.1520691 -0.233007  
-0.093008 0.1336485 -0.01553 -0.0934 -0.067279 0.1098725 -0.007544  
0.0473454 0.0177979 -0.031317 1.5111E-9 0 0 0 0

F1 0.6458627 -0.369546 0.1041262 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
0 0 2.3648E-9 0 0 0 0

F2 0.4843539 0.5154379 -0.018174 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
0 0 -3.92E-10 0 0 0 0

F3 -0.115805 0.0948001 0.5047157 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
0 0 1.068E-10 0 0 0 0

;  
proc iml;  
RESET print;  
USE work.factOut1\_;

```

READ all INTO FactOut1_ VAR {z1 z2 z3 z4 z5 z6 z7 z8 z9 z10 z11 z12 z13
                               z14 z15 z16 z17 z18 z19 z20 z21 z22 z23 z24};

Xz = t(FactOut1_[,1:24]);
XzXzL = Xz*t(Xz);
XzLXz = t(Xz)*Xz;
call svd(UXz,BXz,VXz,Xz);
DBXz = diag(sqrt(Xz));
*Biplot SQRT;
GXz = UXz*DBXz;
HXz = DBXz*t(VXz);
YXz = GXz//t(HXz);

run;

data BYXz_var_obj ;
  input map_VAR_MUNI $ X Y Z;
  cards;
V1 -0.692568 0.4166284 1.1406084 0.1804395 0.8931823 0.048614 0.4439346
0.3473878 -0.10814 -0.058536 0.085528 -0.097842 -0.093591 -0.067375
-0.020226 0.14583260.0958084 0.0478809 -0.155057 -0.052375 -0.042764
0.2177028 -0.143963 0.006422 0 0
V2 0.4432374 1.3329375 -0.285032 -0.150404 -0.040131 -0.314744 0.2019436
0.3587519 -0.128206 0.182496 0.4371479 0.2952499 0.4458496 -0.189272
0.0084452 0.3880819 0.1046158 -0.019386 -0.231633 0.0958866 0.0232004
-0.032957 0.0977418 0.0247869 0 0
V3 -0.705026 0.2150865 1.2313376 0.2556124 0.7920422 -0.011627 0.5455463
0.337235 0.013141 0.0326396 -0.066865 -0.211585 -0.117625 0.0436292
-0.06172 -0.024777 0.09668 -0.028065 0.1658755 0.0396579 0.0537213
-0.215068 0.1378682 0.0114876 0 0
V4 -0.184237 -0.768387 -0.920263 0.1089322 1.1025844 -0.665047 0.4289792
-0.024523 -0.508047 -0.665049 0.2863049 0.3614407 0.1690991 0.1489538

```

0.1571197 -0.081068 0.0177998 -0.017694 0.1366144 0.040642 -0.002311  
0.0041474 -0.017079 0.0060246 0 0  
V5 1.055625 0.0645696 -0.265776 -0.475665 0.7449811 0.0860439 -0.179064  
0.0472716 0.1680561 -0.074137 0.0397109 -0.209348 0.0649646 0.055953  
-0.605977 -0.219387 -0.157308 -0.107916 -0.032062 0.0037378 -0.038573  
0.1914103 0.112967 0.0164416 0 0  
V6 -0.99543 -0.094173 0.1856375 0.3129 0.1876452 0.1183767 -0.654654  
0.4904467 0.9928918 -0.516255 -0.185001 0.4269351 0.293933 0.3766461  
0.0597116 -0.045743 0.0855196 -0.037752 -0.102361 0.0099687 -0.021137  
-0.016461 0.0151631 0.0006951 0 0  
V7 1.0143528 0.3216189 0.0780747 0.3667499 -0.678436 0.239609 0.0949628  
0.4021281 -0.465588 -0.313389 0.0814735 -0.257743 0.1018615 0.747377  
-0.291562 0.192785 0.0186812 0.0711846 0.0909895 0.0119972 -0.023878  
-0.049329 -0.06332 0.031734 0 0  
V8 1.0674062 0.4725804 0.1449824 -0.509885 0.1738519 -0.382551 -0.106941  
0.1165642 0.0349056 -0.022303 -0.124298 -0.215083 -0.282947 0.2483272  
0.3858189 0.0213303 0.2385625 0.1383205 0.0534549 0.0110556 -0.24132  
0.060679 0.0994197 -0.060417 0 0  
V9 1.0212234 0.1401691 0.1281469 0.6163626 0.442731 0.109107 0.0321959  
-0.601105 0.5899243 0.066432 0.1228721 -0.328487 0.6429671 -0.233586  
-0.019965 0.077393 0.0884682 0.1773028 0.2190961 0.0056805 -0.093797  
-0.035353 -0.052591 0.0168987 0 0  
V10 0.859959 -0.146329 -0.157428 -0.86141 0.7141956 -0.144509 -0.849254  
0.3042383 0.1947103 -0.210328 0.3370625 -0.393122 -0.333168 -0.124264  
0.0738575 0.1821738 -0.056053 0.0965492 -0.044715 -0.028285 0.1119296  
-0.14483 -0.07775 0.0347623 0 0  
V11 0.0078563 -1.417646 -0.084001 0.0649861 0.1933061 0.0090712 -0.093538  
0.5518104 -0.241997 0.8394414 0.0328059 -0.243393 0.3533681 0.1739233  
0.1711412 -0.266876 0.1839956 0.113626 -0.10868 -0.059426 0.0275747  
0.012697 0.0292742 0.0547174 0 0

V12 0.9923542 0.7698476 0.1930876 -0.159641 -0.157175 -0.12686 0.3814207  
-0.073001 0.0444764 -0.375346 -0.139409 -0.111246 0.0039797 -0.088192  
-0.039255 -0.587413 0.2905738 0.0197029 -0.178183 0.1525206 0.0989661  
-0.089777 -0.057525 0.002048 0 0

V13 1.0326475 0.4218676 0.0704918 -0.632272 0.3979571 -0.108172 0.3445499  
-0.080646 .0389892 0.2597193 -0.30466 0.4014072 0.2548945 0.2123449  
-0.057435 -0.019986 -0.076946 -0.007387 -0.048361 -0.34225 0.0626198  
-0.102983 -0.044293 -0.062245 0 0

V14 0.734175 0.9442007 -0.200891 0.3105289 -0.120035 -0.820632 -0.097914  
0.6544148 0.1873421 0.2975846 -0.511673 0.0479515 0.0032098 -0.08001  
0.1899794 -0.065886 -0.179379 -0.093066 0.249486 0.0521032 0.1006828  
0.1134875 -0.050849 0.038356 0 0

V15 0.829108 0.3753813 -0.173709 1.1141272 0.4148936 0.0486827 0.064985  
-0.358457 0.318161 0.4735416 0.5638055 0.0407923 -0.317681 0.4025421  
0.2101151 -0.119318 -0.242384 -0.076282 -0.158735 0.0658248 0.0278978  
-0.005154 -0.015364 -0.023106 0 0

V16 0.7755675 -0.274541 -0.128581 1.3057391 0.2438054 -0.423211 -0.33832  
0.0123647 -0.144971 0.0555897 -0.309387 0.4069187 -0.445808 -0.22499  
-0.351062 0.0943174 0.2508684 0.1318726 -0.05895 -0.095984 -0.031238  
-0.056909 0.0289537 0.0206587 0 0

V17 0.8264392 -0.78146 0.3544661 -0.631811 0.4853019 0.3114187 -0.034059  
-0.165908 -0.050544 0.4665952 -0.416049 0.511161 -0.010236 0.1285525  
-0.102939 0.1977041 0.0229847 -0.058755 0.0188572 0.321718 -0.034924  
-0.039808 -0.065278 -0.008565 0 0

V18 1.047763 -0.087927 0.6696952 -0.316284 -0.168099 0.3821853 0.146088  
-0.544922 0.0714711 -0.264372 0.0290833 0.4111232 -0.154355 0.1062627  
0.2947773 0.0301843 0.0446408 0.0028113 0.0384515 -0.097395 0.0827431  
0.095325 0.068394 0.1240554 0 0

V19 -0.337926 0.2776869 -1.420768 -0.0766 0.3270511 0.8674997 0.8879664  
0.2748326 0.4939356 0.0597271 -0.352892 -0.146239 -0.293576 0.0133632

0.0309598 0.128082 0.0728345 0.1079442 -0.01112 -0.005642 0.0012643  
-0.013344 0.0009907 0.0312401 0 0  
V20 -0.971857 0.7705194 -0.247644 -0.222584 0.2837853 -0.017629 -0.281048  
-0.314368 0.0293634 0.4063728 0.2473782 -0.067855 -0.122448 0.2090321  
-0.11488 0.0068868 0.4519562 -0.356374 0.1887018 -0.054951 0.0329394  
0.010827 -0.027988 0.0169371 0 0  
V21 0.967586 -0.750547 0.0387522 0.5337256 0.1401274 0.2070184 0.0077079  
-0.05677 -0.043702 -0.292161 -0.306644 -0.433631 0.1308156 -0.004003  
0.1512732 0.2901364 0.1666042 -0.190037 -0.049876 0.0123271 0.246328  
0.0882797 0.056155 -0.070541 0 0  
V22 0.8410902 -0.773236 0.2428799 -0.16655 -0.532282 0.162647 0.4142131  
0.6396566 0.4778204 0.0610305 0.7312 0.3793524 -0.176818 -0.174284  
-0.131444 -0.052819 0.1357763 0.0402826 0.2098541 -0.007568 0.0632282  
0.0507519 -0.005966 -0.052507 0 0  
V23 -1.017952 0.7424957 -0.061927 -0.153889 0.2235948 0.0614131 -0.339484  
-0.308067 -0.143224 0.1694686 0.0244536 0.1589063 0.0061552 0.2212255  
-0.096037 -0.030154 0.0459194 0.4497406 0.1012778 0.0707292 0.2273902  
0.0993623 0.0386513 -0.035884 0 0  
V24 0.5444837 0.7915465 -0.077073 0.3927548 0.4301952 1.3206754 -0.584688  
0.4229898 -0.625888 -0.088167 0.039792 0.25013 0.1021727 -0.246778 0.192043  
-0.1903 -0.022538 -0.053014 0.101174 -0.020857 -0.041857 -0.01203 0.0022159  
-0.022132 0 0  
01 -0.979627 0.7876227 -0.227114 -0.224631 0.6182168 0.5322924 -0.902362  
-0.023954 -0.129501 -0.182199 0.2018287 0.2124588 -0.035949 -0.295059  
0.1519753 0.1006947 0.1832548 0.0404654 -0.20182 -0.057084 -0.002748  
0.0842448 -0.083309 0.0563811 1.9401E-9 0  
02 -0.493499 -0.155206 -0.404179 -0.529248 0.4545668 -0.002635 0.0354267  
-0.037959 0.0447122 -0.201175 0.2478029 -0.258447 -0.502012 -0.31204  
0.0107509 0.1156407 -0.270555 -0.055302 0.1722951 -0.012718 0.196736  
-0.15984 -0.044579 -0.041668 1.9674E-9 0

03 -0.754406 -0.194731 -0.231972 0.005272 0.0038611 0.0605881 -0.146795  
0.2219013 0.5318994 -0.508463 0.0950639 0.4370133 0.160997 -0.046385  
0.252351 -0.241012 -0.145854 -0.261427 0.057794 0.0728834 -0.149237  
-0.074554 0.0169965 -0.006994 -2.42E-10 0  
04 -0.314969 -0.601577 -0.220807 0.5666614 -0.366999 -0.176791 0.0945273  
0.2590169 0.0888832 0.2326422 -0.732647 0.0659177 -0.27406 -0.448793  
-0.186677 0.1700328 0.1215694 0.0272435 0.023853 -0.211349 -0.049963  
-0.022046 -0.039432 0.0133115 1.861E-10 0  
05 -0.482704 -0.362219 -0.328704 0.3746806 -0.149037 -0.114561 -0.43903  
0.3645751 0.5222445 -0.418943 -0.387522 0.2623269 0.1567615 0.4140485  
-0.203433 0.0382838 -0.072161 0.2204452 -0.066824 0.0356803 0.0805933  
0.029881 -0.049796 -0.070368 2.6971E-9 0  
06 -0.837079 -0.049673 -0.314183 0.1113902 -0.334108 -0.166275 0.1529743  
-0.454167 -0.139722 0.0429124 -0.113342 -0.023378 0.0651311 -0.173437  
0.0221121 -0.083777 -0.225867 0.050328 0.3117265 0.0926086 -0.076133  
0.2736949 -0.035866 0.0347612 2.2938E-9 0  
07 -1.586809 0.51105 1.78368 0.4508218 0.9312483 0.0237927 0.4916021  
0.3527881 -0.066963 -0.011618 -0.050742 -0.107179 0.0006728 0.0736565  
-0.062272 0.0217225 -0.021226 0.0032931 0.0290895 -0.002955 0.008477  
0.0440852 0.0152505 -0.017723 1.2607E-9 0  
08 -1.171448 -0.005642 -0.012824 0.1061633 -0.18579 -0.153457 0.2871778  
-0.451367 0.3023806 0.2980289 -0.227794 0.011063 -0.034713 0.1155931  
0.0071018 -0.115672 0.0577313 0.068276 -0.074642 0.1208036 -0.033897  
-0.248069 0.0217941 0.0973941 2.7778E-9 0  
09 0.0385102 -1.060244 -0.474448 -0.35581 0.313375 0.2036393 0.0895891  
0.9187835 -0.281504 0.4949697 0.0193325 -0.403996 0.1283205 -0.024943  
0.1515151 -0.282699 -0.058648 0.1832022 -0.097757 0.1399199 0.0342813  
0.0304032 -0.017854 0.0217541 .882E-10 0  
010 0.2214523 -0.692099 -0.227534 -0.598155 0.0991518 0.105999 0.3069666  
0.4747851 0.1255012 -0.178845 0.1005423 -0.117605 -0.070426 0.5172466



0.0576058 0.3104125 0.2577845 -0.229058 0.1772054 -0.101552 -0.030523  
0.0500623 -0.011973 0.0565628 2.1674E-9 0  
011 0.2482817 -0.689219 1.245794 -0.288247 -1.005974 -0.120862 -1.020659  
0.0856021 0.0460769 0.2960783 0.3796797 -0.051642 -0.218726 0.0264233  
0.0782663 -0.008278 -0.041715 0.0955114 0.0938843 -0.049057 -0.072918  
-0.035423 0.0315714 -0.005612 1.6722E-9 0  
012 -0.701748 0.4269594 -0.432144 -0.106974 -0.003773 0.3100445 -0.496259  
-0.206673 -0.231807 0.11334 -0.210785 0.0098983 0.1017112 0.1395622  
-0.148777 -0.034025 0.1615212 -0.013197 0.1474058 0.0209535 0.1887311  
0.0203726 0.2100691 -0.001162 4.304E-10 0  
013 -0.850054 -0.409212 -0.033247 0.0324532 -0.408989 -0.575391 0.1356281  
-0.410154 0.2338107 0.3277299 0.2747581 -0.298124 0.3226467 -0.058524  
0.0460063 -0.235588 0.2091544 -0.260557 -0.157694 -0.105303 0.1346174  
0.0682283 -0.059019 -0.05204 1.6265E-9 0  
014 0.2022127 1.0128206 -0.26164 0.2369676 -0.272892 0.0888161 0.0472006  
-0.03457 -0.086661 0.5846233 0.1932548 0.2437453 -0.543526 0.3688426  
-0.223469 0.0505203 0.0339367 -0.109399 -0.042069 0.1917313 -0.020592  
0.0323101 -0.099821 -0.024841 4.004E-10 0  
015 1.0241038 1.4723371 -0.128993 0.0164726 0.0681329 -0.702239 0.026583  
0.4246642 0.1930593 0.5410164 0.0524488 0.3144784 0.4337487 -0.031459  
0.2425861 0.1678713 -0.192816 0.0315142 0.0770836 -0.111736 0.0758718  
-0.037133 0.0046837 0.0224452 1.6292E-9 0  
016 0.5359919 -0.088153 0.3030679 -0.916673 0.0822421 -0.379707 0.2309791  
-0.477963 -0.243982 -0.455031 0.0734734 0.1873053 0.0724087 0.253893  
-0.25888 -0.211108 -0.077965 0.2089015 -0.031006 -0.133433 0.0615569  
-0.007083 -0.073088 0.0497503 -7.13E-11 0  
017 0.7351063 0.871196 -0.053236 -0.553379 0.0808205 -0.619388 -0.193807  
0.3462783 -0.477998 -0.278395 -0.422642 -0.193399 -0.219699 -0.085545  
-0.187482 -0.332318 0.0032144 -0.194072 -0.091441 -0.008747 -0.128182  
-0.000414 0.0540607 -0.020028 2.7606E-9 0

018 0.7879182 0.4172565 0.3941929 0.0103596 -0.630265 -0.382631 0.0967189  
-0.04231 -0.009599 -0.623975 -0.195817 -0.409169 0.1054475 -0.209302  
0.24629 0.3063133 0.1963825 0.0493501 -0.036183 0.2594726 0.0733357  
0.0000102 -0.023508 0.0051872 7.379E-10 0  
019 -0.742861 -0.020385 -0.60761 -0.419452 0.3449812 -0.40927 0.3005612  
-0.331002 -0.218762 0.096264 0.2696813 -0.038091 0.0226453 0.0521381  
0.1637249 0.2556847 0.123185 0.2518628 -0.037418 -0.037607 -0.229256  
-0.033955 0.0640492 -0.084256 1.2619E-9 0  
020 -0.041697 -0.130612 -0.146782 0.2108794 -0.549945 0.2157295 0.3808922  
0.0401473 0.1359721 -0.175203 0.2879561 -0.112639 -0.147872 -0.011509  
-0.122849 0.2434336 -0.361524 -0.059387 -0.344385 -0.056581 0.0082819  
0.1048027 0.1225962 0.0348873 1.4916E-9 0  
021 0.8976767 -1.262859 -0.040143 1.1964687 0.6364817 -0.559729 -0.180386  
-0.124203 -0.689387 -0.11131 0.3315107 0.4336303 0.0840775 -0.069109  
-0.107438 0.0792658 0.0322192 -0.09068 0.0013209 0.0782401 0.0289983  
-0.041048 0.0171324 0.0163154 1.8493E-9 0  
022 1.2261587 0.3166123 -0.192214 0.7453841 0.6998952 0.3123928 -0.424722  
-0.353758 0.5111363 0.0131943 0.0945421 -0.825267 0.1973889 0.0876112  
-0.230945 -0.024426 -0.061033 0.0051866 0.064118 -0.043898 -0.11424  
-0.031857 -0.01901 0.0102174 1.0477E-9 0  
023 1.3121164 0.1201337 0.0205704 0.6063005 0.1408075 0.2847574 0.4234821  
-0.162084 0.3541192 -0.142124 0.031456 0.1791981 -0.512073 0.0228991  
0.4431753 -0.349672 .1670287 0.1522939 0.0123578 -0.088056 0.0599362  
0.071659 0.0502899 -0.010579 1.8813E-9 0  
024 0.2949377 0.4776039 -0.038999 0.0525766 -0.700632 0.809493 0.5510248  
0.3597649 -0.086422 -0.162613 0.4957613 0.1837141 0.3269271 -0.343267  
-0.388937 -0.111322 0.1764015 0.0761709 0.10653 -0.004042 -0.028539  
-0.073181 -0.008656 -0.019219 2.1568E-9 0  
025 0.1247341 0.0339197 0.0999973 0.2843538 -0.438989 0.910055 0.0474288  
-0.302735 -0.828359 0.0027118 -0.42311 -0.077899 0.2281039 0.2704653

98

```
0.3407066 0.0363721 -0.178488 -0.097565 -0.026146 -0.098047 -0.007643
-0.092495 -0.060381 -0.033163 1.6318E-9 0
026 1.3076994 -0.725681 0.5294705 -1.014635 0.5736125 0.5053364 0.1052562
-0.435409 0.4008712 0.4063818 -0.384691 0.376085 0.1520691 -0.233007
-0.093008 0.1336485 -0.01553 -0.0934 -0.067279 0.1098725 -0.007544
0.0473454 0.0177979 -0.031317 1.8414E-9 0
```

```
;
```

```
proc plot data=BYXz_var_obj;
plot X*Y = '.'$ map_VAR_MUNI;
quit;
run;
```

```
/*
```

Para obter o biplot em 3D no software SAS:

Solution/ analysis/ interative data analysis

Work/ escolha do arquivo/ open

Analyse/ rotating plot

Coloque Z em Z, Y em Y, X em X e nome do arquivo em lable

Ok

```
*/
```

## ANEXO E - Tabela contendo as cargas fatoriais antes da rotação varimax

Variáveis	Fator 1	Fator 2	Fator 3	Variáveis	Fator 1	Fator 2	Fator 3
V1	-0,57040	0,26980	0,59397	V13	<u>0,85049</u>	0,27320	0,03671
V2	0,36505	<u>0,86320</u>	-0,14843	V14	0,60467	0,61145	-0,10461
V3	-0,58066	0,13929	0,64121	V15	0,68286	0,24309	-0,09046
V4	-0,15174	-0,9760	-0,47922	V16	0,63876	-0,17779	-0,06696
V5	0,86942	0,04181	-0,13840	V17	0,68066	-0,50606	0,18459
V6	<u>-0,81984</u>	-0,06099	0,09667	V18	<u>0,86294</u>	-0,05694	0,34874
V7	<u>0,83542</u>	0,20828	0,04066	V19	-0,27832	0,17983	<u>-0,73986</u>
V8	<u>0,87912</u>	0,30604	0,07550	V20	<u>-0,80042</u>	0,49898	-0,12896
V9	<u>0,84108</u>	0,09077	0,06673	V21	<u>0,79691</u>	-0,48605	0,02018
V10	<u>0,70826</u>	-0,09476	-0,08198	V22	0,69272	-0,50074	0,12648
V11	0,00647	<u>-0,91805</u>	-0,04374	V23	<u>-0,83839</u>	0,48083	-0,03225
V12	<u>0,81731</u>	0,49854	0,10055	V24	0,44844	0,51260	-0,04014

## ANEXO F - Tabela contendo as correlações entre as variáveis

Variáveis	V1	V2	V3	V4	V5	V6	V7	V8
V1	1,0000	-0,0493	<u>0,9808</u>	-0,0656	-0,4428	0,5237	-0,4862	-0,3648
V2	0,0493	1,0000	-0,1834	-0,3617	0,3672	-0,4016	0,4742	0,5946
V3	0,9808	-0,1834	1,0000	-0,0518	-0,4884	0,5431	-0,5026	-0,4182
V4	-0,0656	-0,3617	-0,0518	1,0000	0,0584	0,0986	-0,3890	-0,2615
V5	-0,4428	0,3672	-0,4884	0,0584	1,0000	-0,6946	0,5752	<u>0,8504</u>
V6	0,5237	-0,4016	0,5431	0,0986	-0,6946	1,0000	-0,7048	-0,7549
V7	-0,4862	0,4742	-0,5026	-0,3890	0,5752	-0,7048	1,0000	0,7201
V8	-0,3648	0,5946	-0,4182	-0,2615	0,8504	-0,7549	0,7201	1,0000
V9	-0,3252	0,3239	-0,3326	-0,1242	0,7431	-0,5990	0,6543	0,6846
V10	-0,3252	0,3239	-0,3326	-0,1242	0,7431	-0,5990	0,6543	0,6846
V11	-0,2181	-0,7444	-0,0992	0,4744	-0,0030	0,0319	-0,1876	-0,2689
V12	-0,2911	0,7149	-0,3506	-0,4072	0,7063	-0,7318	<u>0,8011</u>	0,8985
V13	-0,3270	0,5820	-0,3801	-0,1738	0,8759	-0,7672	0,6302	0,9266
V14	-0,2466	0,8101	-0,3082	-0,3219	0,4954	-0,4743	0,6433	0,7384
V15	-0,2537	0,4200	-0,2787	-0,0790	0,5489	-0,4965	0,6215	0,5335
V16	-0,3649	0,0434	-0,3361	0,1521	0,4341	-0,3908	0,5446	0,3705
V17	-0,3592	-0,2141	-0,3142	0,1038	0,6988	-0,5443	0,3179	0,5319
V18	-0,3641	0,1852	-0,3585	-0,3273	0,6957	-0,7072	0,7058	0,7679
V19	-0,0919	0,1241	-0,1394	0,2923	-0,0672	0,1518	-0,2496	-0,2765
V20	0,5300	0,1515	0,4548	-0,0395	-0,5732	0,5923	-0,6607	-0,5157
V21	-0,5147	-0,1873	-0,4518	0,1426	0,6333	-0,5691	0,6201	0,4854
V22	-0,5186	-0,1251	-0,4472	-0,0603	0,4930	-0,5109	0,5568	0,4437
V23	0,5909	0,0967	0,5203	-0,0680	-0,6425	0,6508	-0,6605	-0,5678
V24	0,5909	0,0967	0,5203	-0,0680	-0,6425	0,6508	-0,6605	-0,5678

## ANEXO F - Tabela contendo as correlações entre as variáveis

(continuação)

Variáveis	V9	V10	V11	V12	V13	V14	V15	V16
V1	-0,3252	-0,4058	-0,2181	-0,2911	-0,3270	-0,2466	-0,2537	-0,3649
V2	0,3239	0,2024	-0,7444	0,7149	0,5820	<u>0,8101</u>	0,4200	0,0434
V3	-0,3326	-0,4516	-0,0992	-0,3506	-0,3801	-0,3082	-0,2787	-0,3361
V4	-0,1242	0,0950	0,4744	-0,4072	-0,1738	-0,3219	-0,0790	0,1521
V5	0,7431	<u>0,8862</u>	-0,0030	0,7063	<u>0,8759</u>	0,4954	0,5489	0,4341
V6	-0,5990	-0,5134	0,0319	-0,7318	-0,7672	-0,4743	-0,4965	-0,3908
V7	0,6543	0,3705	-0,1876	0,8011	0,6302	0,6433	0,6215	0,5446
V8	0,6846	0,7694	-0,2689	<u>0,8985</u>	<u>0,9266</u>	0,7384	0,5335	0,3705
V9	1,0000	0,5024	-0,0862	0,7044	0,6862	0,5352	<u>0,8521</u>	0,7009
V10	0,5024	1,0000	0,1215	0,4887	0,7017	0,3306	0,2812	0,2623
V11	-0,0862	0,1215	1,0000	-0,5047	-0,2296	-0,4866	-0,1712	0,1941
V12	0,7044	0,4887	-0,5047	1,0000	<u>0,8582</u>	0,7811	0,5887	0,3518
V13	0,6862	0,7017	-0,2296	0,8582	1,0000	0,6426	0,5065	0,3050
V14	0,5352	0,3306	-0,4866	0,7811	0,6426	1,0000	0,6075	0,4620
V15	0,8521	0,2812	-0,1712	0,5887	0,5065	0,6075	1,0000	0,7746
V16	0,7009	0,2623	0,1941	0,3518	0,3050	0,4620	0,7746	1,0000
V17	0,4867	0,6858	0,4913	0,3003	0,6326	-0,0081	0,2041	0,3376
V18	0,7111	0,5779	-0,0465	0,7415	0,7779	0,3195	0,4663	0,3934
V19	-0,2214	-0,1970	-0,1263	-0,1878	-0,1407	-0,0861	-0,0570	-0,2736
V20	-0,6170	-0,4809	-0,4296	-0,4499	-0,4948	-0,2121	-0,4008	-0,6378
V21	0,7452	0,5145	0,4466	0,3947	0,4545	0,1882	0,5598	0,7547
V22	0,4460	0,4660	0,4524	0,3602	0,4384	0,1067	0,2722	0,3980
V23	-0,6550	-0,5476	-0,4359	-0,4783	-0,5411	-0,2590	-0,4557	-0,6377
V24	0,4734	0,2945	-0,3905	0,5192	0,4352	0,4445	0,5449	0,3007

## ANEXO F - Tabela contendo as correlações entre as variáveis

(conclusão)

Variáveis	V17	V18	V19	V20	V21	V22	V23	V24
V1	-0,3592	-0,3641	-0,0919	0,5300	-0,5147	-0,5186	0,5909	-0,0358
V2	-0,2141	0,1852	0,1241	0,1515	-0,1873	-0,1251	0,0967	0,5381
V3	-0,3142	-0,3585	-0,1394	0,4548	-0,4518	-0,4472	0,5203	-0,1489
V4	0,1038	-0,3273	0,2923	-0,0395	0,1426	-0,0603	-0,0680	-0,3243
V5	0,6988	0,6957	-0,0672	-0,5732	0,6333	0,4930	-0,6425	0,4630
V6	-0,5443	-0,7072	0,1518	0,5923	-0,5691	-0,5109	0,6508	-0,3018
V7	0,3179	0,7058	-0,2496	-0,6607	0,6201	0,5568	-0,6605	0,5388
V8	0,5319	0,7679	-0,2765	-0,5157	0,4854	0,4437	-0,5678	0,4290
V9	0,4867	0,7111	-0,2214	-0,6170	0,7452	0,4460	-0,6550	0,4734
V10	0,6858	0,5779	-0,1970	-0,4809	0,5145	0,4660	-0,5476	0,2945
V11	0,4913	-0,0465	-0,1263	-0,4296	0,4466	0,4524	-0,4359	-0,3905
V12	0,3003	0,7415	-0,1878	-0,4499	0,3947	0,3602	-0,4783	0,5192
V13	0,6326	0,7779	-0,1407	-0,4948	0,4545	0,4384	-0,5411	0,4352
V14	-0,0081	0,3195	-0,0861	-0,2121	0,1882	0,1067	-0,2590	0,4445
V15	0,2041	0,4663	-0,0570	-0,4008	0,5598	0,2722	-0,4557	0,5449
V16	0,3376	0,3934	-0,2736	-0,6378	0,7547	0,3980	-0,6377	0,3007
V17	1,0000	0,7484	-0,3214	-0,7340	0,7129	0,6963	-0,7402	0,0957
V18	0,7484	1,0000	-0,4551	-0,7534	0,6847	0,6968	-0,7396	0,3476
V19	-0,3214	-0,4551	1,0000	0,3815	-0,2830	-0,3050	0,3066	0,1187
V20	-0,7340	-0,7534	0,3815	1,0000	<u>-0,9189</u>	<u>-0,8710</u>	<u>0,9734</u>	-0,0994
V21	0,7129	0,6847	-0,2830	-0,9189	1,0000	0,7179	<u>-0,9254</u>	0,2170
V22	0,6963	0,6968	-0,3050	-0,8710	0,7179	1,0000	<u>-0,8937</u>	-0,0101
V23	-0,7402	-0,7396	0,3066	0,9734	-0,9254	-0,8937	1,0000	-0,0847
V24	0,0957	0,3476	0,1187	-0,0994	0,2170	-0,0101	-0,0847	1,0000