

O USO DA ANÁLISE FATORIAL NA DESCRIÇÃO E IDENTIFICAÇÃO DOS PERFIS CARACTERÍSTICOS DE MUNICÍPIOS DE MINAS GERAIS

Naje Clécio Nunes da SILVA¹
Wederson Leandro FERREIRA¹
Marcelo Ângelo CIRILLO²
João Domingos SCALON²

- **RESUMO:** A análise fatorial é uma técnica estatística multivariada utilizada para identificar um número reduzido de fatores latentes a partir de um número de variáveis observadas e correlacionadas. Neste trabalho foi utilizada a análise fatorial para determinar os perfis de atividades predominantes em trinta municípios do estado de Minas Gerais. Para o estudo pretendido, foram aferidas oito variáveis com o objetivo de subsidiar e apoiar o planejamento municipal e regional destes municípios. Utilizando os métodos de eixos principais e rotação ortogonal foi possível caracterizar os municípios através de três fatores (econômico, fundiário e agropecuário). Os resultados também evidenciaram que houve heterogeneidade e discrepância entre os municípios avaliados. Este cenário sugere que existe a necessidade de adotar ações para o maior desenvolvimento destes municípios, propiciando assim, um ambiente favorável à reversão deste quadro.
- **PALAVRAS-CHAVE:** Análise multivariada; economia; agropecuária; estrutura fatorial.

1 Introdução

O estado de Minas Gerais possui 853 municípios sendo a terceira maior economia entre as 27 unidades da Federação Brasileira, possuindo no ano de 2009 um PIB (Produto Interno Bruto) de R\$ 287,1 bilhões equivalente a 8,9% do total do Brasil, cujo PIB foi de R\$ 3,2 trilhões (IBGE, 2009). Nesta mesma década, houve elevação de seu valor agregado à produção nacional, progredindo de 8,4% em 1999 para 9,4% em 2010. Referindo-se as exportações brasileiras, Minas Gerais, após queda nos anos 90, aumentou de forma consistente sua participação a partir do ano de 2003, que era de 10,2%, passando a ser de 15,5% em 2010, fato que lhe conferiu a posição de segundo maior exportador do Brasil (PMDI, 2011). A evolução desses índices deve-se em grande parte ao aspecto do estado ser detentor de uma economia muito diversificada.

¹ Universidade Federal de Lavras – UFLA, Programa de Pós-graduação em Estatística e Experimentação Agropecuária - PPEEA, Departamento de Ciências Exatas - DEX, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: naje.silva@ifac.edu.br; wedeleian@gmail.com

² Universidade Federal de Lavras – UFLA, Departamento de Ciências Exatas - DEX, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: macufta@dex.ufla.br; scalon@dex.ufla.br

A diversidade da economia Mineira, segundo Perobelli et al. (1999), apresenta potencialidades expressivas, destacando-se a base minero-metalúrgica, a base universitária, o desenvolvimento de sua malha de transporte, a ampliação de seu potencial exportador, o perfil da mão-de-obra mineira, dentre outros fatores.

Na área industrial o estado também apresenta algumas características superiores a outras unidades da federação. Por exemplo, o estado é o maior produtor e exportador mundial de ferro-nióbio, o maior produtor e exportador brasileiro de ferro gusa, o maior produtor brasileiro de cimento e o segundo polo automotivo do país, dentre outros. Além disso, o estado abriga o maior e mais completo parque siderúrgico do Brasil. As plantas industriais estão instaladas em municípios como Juiz de Fora, Ouro Branco, João Monlevade e Ipatinga, que, juntos, produzem todos os tipos de aços requisitados pela indústria: semiacabados, laminados planos, longos, relaminados, trefilados e perfilados (SEDE, 2013).

Outra área em que o estado é referência é na agropecuária. Dentre os estados do sudeste, Minas Gerais é o maior produtor agropecuário. Estima-se que 60% do produto agropecuário das regiões Sul, Sudeste e Centro-Oeste são explicados pela produção dos estados de São Paulo, Rio Grande do Sul e Minas Gerais, podendo-se inferir que Minas Gerais é uma das regiões mais produtivas na agropecuária brasileira, tornando-a objeto de vários estudos relacionados ao setor no Brasil (Clemente e Gomes, 2011).

Minas Gerais também se destaca na produção de leite, sendo o maior produtor nacional, e esta atividade é considerada como uma das seis atividades agropecuárias mais importantes do país (NASCIMENTO et al, 2012). Nesta atividade, destacam-se as regiões da Zona da Mata e o Sul de Minas, com promissor crescimento das regiões do Triângulo Mineiro e do Alto Paranaíba, que têm se intensificado, garantindo a essas regiões posição de liderança neste setor.

A geração de renda no Estado de Minas Gerais tem como uma de suas características principais o alto índice de concentração regional. Por exemplo, a região mais próspera e populosa é a Central que responde por quase metade (46,6%) do PIB mineiro; por outro lado, as regiões historicamente mais deprimidas são Norte, Jequitinhonha/Mucuri e Rio Doce, que juntas totalizam apenas 12,2% (PMDI, 2011).

Ainda, segundo PMDI (2011), "acentuadas disparidades regionais da economia mineira são visíveis também quando se considera o PIB per capita de cada uma delas. No Triângulo Mineiro, região que concentra o mais elevado PIB per capita do estado (R\$ 21 mil), o indicador é mais de quatro vezes superior ao registrado na área mais pobre, a região do Jequitinhonha/Mucuri (R\$ 5,2 mil)".

Diante das complexas disparidades econômicas que ocorrem no Estado de Minas Gerais, é desejável que o perfil das atividades exercidas por seus municípios sejam estudadas e completamente caracterizadas. Isto subsidiará as adoções de ações governamentais no intuito de amenizar tais distorções. Existem municípios que possuem essa caracterização bem definida, mas outros municípios geralmente de pequeno porte, não estão completamente caracterizados quanto aos perfis das atividades existentes, impondo, desta forma, limitações na adoção de políticas públicas.

Conforme apresenta PMDI (2011), "com uma estratégia regionalizada, com soluções customizadas, direcionadas às necessidades específicas de cada local, é possível reduzir as disparidades do território, colocando o Estado por inteiro no caminho do desenvolvimento. A estratégia de regionalização consiste, portanto, em focalizar ações nos territórios, de modo a potencializar suas vantagens comparativas e compensar as

carências territoriais, minimizando as assimetrias regionais. Em outras palavras, é preciso propor soluções e estratégias condizentes para o crescimento econômico e social das diversas regiões mineiras".

O estudo dos fenômenos industriais, fundiários, pecuários e agropecuários, dentre outros, envolve a análise simultânea de várias variáveis, e então é desejável a obtenção de um entendimento geral da relação existente entre tais variáveis, o que pode ser obtido através de técnicas estatísticas de análise multivariada.

A estatística multivariada, segundo Johnson & Wichern (2007), permite o estudo de fenômenos complexos, pois realiza o tratamento de diversas variáveis simultaneamente, mesmo quando não se conhece o modelo teórico que as relaciona. De acordo com Bakke et al. (2008), a análise fatorial, uma das técnicas da estatística multivariada, pode ser aplicada quando há uma grande massa de dados e busca-se a sua sumarização com fatores latentes, não observáveis, que são definidos através de comportamento semelhante entre as variáveis.

Pela diversidade de atividades apresentadas pelo estado de Minas Gerais, o objetivo deste trabalho é estabelecer os perfis característicos predominantes em 30 municípios por meio do método da análise fatorial, buscando subsidiar possíveis planejamentos públicos, ou opções empresariais, que norteiem a diversificação de atividades em municípios que possuem algum perfil específico ou apoiar ações naqueles em que as atividades não estão caracterizadas.

2 Material e métodos

Os dados utilizados neste trabalho foram oriundos do Zoneamento Ecológico Econômico (ZEE) de 30 municípios do estado de Minas Gerais (SCOLFORO et al., 2008a) (SCOLFORO et al., 2008b) (SCOLFORO et al., 2008c). Os municípios avaliados foram: Araxá, Arcos, Ataléia, Augusto de Lima, Barbacena, Barroso, Cachoeira de Minas, Campo Belo, Douradoquara, Durandé, Elói Mendes, Extrema, Formiga, Formoso, Ipaba, Jaboticatubas, Jordânia, Lavras, Mata Verde, Morro da Garça, Nova Resende, Ouro Branco, Pouso Alegre, Rubim, São Vicente de Minas, Tiradentes, Tiros, Três Corações, Urucânia e Urucuia. Nesses locais, foram aferidas oito variáveis: Índice do valor adicionado agropecuário (IVAA); Índice do valor adicionado industrial (IVAI); Índice do valor adicionado de serviços (IVAS); Índice do produto interno bruto (PIB); Índice de concentração fundiária invertido (ICFI); Índice de Gini da concentração fundiária (GINI); Índice de agricultores familiares (IAF) e Índice de emprego formal (IEF). Informações adicionais sobre a formulação dos índices adimensionais das 8 variáveis, podem ser consultadas nas referências citadas anteriormente.

O procedimento estatístico utilizado para atingir os objetivos propostos neste trabalho foi a análise fatorial que essencialmente busca descrever a estrutura de covariâncias, ou de correlações, entre as variáveis originais em termos de um número menor de variáveis latentes (variáveis inobserváveis) denominadas de fatores. Desta forma, ao sumarizar os dados, há a captação das dimensões latentes que representará o conjunto de dados em um número menor de conceitos do que as variáveis individuais originais (Johnson & Wichern 2007; BAKKE et al. 2008). De acordo com Macciotta et al. (2012), é possível ainda saber o quanto cada fator está associado a cada variável original e o quanto o conjunto de fatores explica a variabilidade total dos dados originais.

Segundo Moori e Zilber (2003), a análise fatorial pode ser utilizada com alguns objetivos particulares, como, por exemplo: descobrir a estrutura de um conjunto de medições; reduzir a grande massa de dados da pesquisa em quantidade gerenciável; reagrupar as variáveis originais possivelmente dependentes entre si, em tese, em fatores latentes, não observáveis, independentes.

O modelo fatorial adotado para uma variável X_i , observável, com média μ_i pode ser representado da seguinte forma (MARDIA et al. 2006; Johnson & Wichern 2007):

$$X_i - \mu_i = l_{i1}F_1 + l_{i2}F_2 + \dots + l_{im}F_m + \varepsilon_i, \quad (1)$$

onde: $i = 1, 2, \dots, p$ e $m \leq p$, sendo p o número de variáveis originais observáveis; O coeficiente l_{ij} é chamado de carga fatorial da i -ésima variável sobre o j -ésimo fator comum, sendo $j = 1, 2, \dots, m$; F_1, F_2, \dots, F_m são denominados de fatores comuns, variáveis aleatórias inobserváveis e ε_i são os erros aleatórios que estão associados somente a i -ésima resposta X_i , respectivamente. Este modelo postula que as variáveis originais se relacionam de forma linearmente dependente com alguns dos fatores comuns e com os p erros.

As cargas fatoriais são as correlações de cada variável com um determinado fator, ou seja, elas indicam o grau de correspondência entre a variável e o fator, sendo as cargas numericamente maiores as responsáveis pela denominação que o fator receberá. O próprio modelo já explicita que o número de fatores deve ser no máximo igual ao número de variáveis observáveis, ressaltando que conforme o número de fatores extraídos se aproxima do número total de variáveis, menos vantajosa torna-se a análise fatorial.

Para a condução da análise fatorial, diversos autores (PEROBELLI et al., 1999; Moori e Zilber, 2003; BAKKE et al., 2008) recomendam seguir um roteiro conforme descrito a seguir:

Em geral, o primeiro passo consiste no exame das correlações entre as variáveis observáveis, procedimento realizado pela obtenção da matriz de correlações. Por meio dessa matriz é possível identificar subconjuntos de variáveis que estão muito correlacionadas entre si no interior de cada subconjunto, mas pouco associados a variáveis de outros subconjuntos. O uso da matriz de correlações, em detrimento da matriz de covariâncias, procura minimizar a influência da magnitude das distintas unidades em que as variáveis foram aferidas (MARDIA et al., 2006). Desta forma, se for preferido utilizar a matriz de covariâncias e sendo as variáveis adimensionais, deve-se realizar a padronização das mesmas, pois este procedimento neutraliza o efeito das distintas unidades, colocando todas elas na mesma escala.

Sequencialmente, faz-se o teste de esfericidade de Bartlett que é utilizado para examinar a hipótese nula de que as variáveis não sejam correlacionadas na população, ou seja, a matriz de correlações da população é uma matriz identidade, por conseguinte, o prosseguimento da análise torna-se adequada se a hipótese nula for rejeitada.

No intuito de determinar o número de fatores a serem utilizados na análise fatorial é necessário calcular os autovalores da matriz de correlações, de tal forma que mantenha um conjunto de vetores independentes, não correlacionados, que expliquem o máximo da variabilidade dos dados. Os autovalores são números que refletem a importância do fator, sendo que divididos pela soma de todos os autovalores indicam a proporção da variabilidade total dos dados que é explicada pelo fator. Mingoti (2005) argumenta que o

número de fatores a serem retidos deve refletir um valor superior a 70% da variabilidade original dos dados, entretanto, esta escolha não é um procedimento trivial³.

Outras medidas que podem ser obtidas paralelamente são as comunalidades, que são proporções das variâncias (ou correlações) de cada variável explicada pelos fatores comuns. Também são interpretadas como índices atribuídos às variáveis originais que expressam, em termos percentuais, o quanto da variabilidade de cada variável é explicada pelo modelo adotado. Desta forma, quanto mais alto forem esses indicadores, melhor será o indício do adequado ajuste do modelo fatorial.

Determinado o número de fatores, calculam-se suas cargas fatoriais que são as correlações entre as variáveis originais e os fatores, sendo um aspecto importante da análise, pois quanto maior a carga fatorial maior será a correlação com determinado fator.

Para a extração dos fatores pode ser utilizado o método de componentes principais (PEROBELLI et al., 1999). Este método baseia-se na decomposição da matriz de covariância ou correlação e não exige pressuposição da distribuição dos dados, tornando-se mais flexível, algo que não ocorre com o método de máxima verossimilhança. A diferença básica dos métodos é que o método de componentes principais utiliza toda a variância (compartilhada e específica) e o erro, constituindo a chamada abordagem exploratória. Já o método da máxima verossimilhança, que exige normalidade multivariada, procura estimar parâmetros que maximizem a probabilidade de observar resultados iguais, caso os dados sejam novamente coletados na mesma população (BROWN, 2006).

Quando não é clara a interpretação dos fatores extraídos utilizando-se as cargas fatoriais, ou seja, existindo mais de um fator e a contribuição das variáveis para cada um deles não é suficientemente clara, um procedimento adequado que pode ser utilizado é a rotação dos eixos coordenados (Johnson & Wichern, 2007).

Existem vários métodos de rotação, sendo que neste trabalho serão considerados os métodos Varimax e o Promax, que procuram dar aos fatores maior potencial de interpretabilidade, ou seja, torna a solução fatorial mais simples e pragmaticamente mais significativa (MARDIA et al., 2006; Johnson & Wichern, 2007).

O método Varimax é um método de rotação ortogonal, sendo o mais comumente utilizado dentre os métodos ortogonais, que procura minimizar o número de variáveis que apresentam altas cargas em cada fator. Já o método Promax é um método de rotação oblíquo em que o pressuposto de independência entre os fatores é retirado, permitindo que os mesmos rodem livremente de maneira a simplificarem a sua interpretação (HAIR et al., 2006).

Para selecionar um dos métodos, será utilizado a Raiz Quadrada do Erro Quadrático Médio (REQM), que avalia a adequabilidade do modelo, conforme Mingoti (2005). Essa medida apresenta um índice de discrepância entre a matriz de covariância amostral e a matriz de covariância obtida através do modelo fatorial, apresentando uma amplitude compreendida entre 0 e 1, sendo que quanto mais próximo de zero melhor o ajuste do modelo.

A análise de ajuste do modelo pode ser realizada pela interpretação dos resíduos que são as diferenças entre termos de covariâncias observadas e termos de covariâncias ajustadas. Quanto menores os resíduos melhor o ajuste, pois são índices interpretados

³ Para outros critérios sobre o número de fatores ver Hair et al. (2006) e Mardia et al. (2006).

como as covariâncias residuais não explicadas pelo modelo (ULLMAN, 2006). Segundo Hair et al. (2006), estes são dependentes do real intervalo da escala de medição, porém quando as variáveis originais são adimensionais utilizam-se os resíduos padronizados, porque são independentes das escalas de medição. A obtenção de seus valores é simplesmente os valores dos resíduos originais divididos pelo erro padrão do respectivo resíduo.

Ainda segundo Hair et al. (2006) os resíduos padronizados podem ser negativos ou positivos, em que valores inferiores a $|2,5|$ indicam ausência de problemas no ajuste do modelo, valores entre $|2,5|$ e $|4,0|$ merecem atenção, mas podem não sugerir mudanças no modelo se outros problemas no ajuste não forem diagnosticados. Já se forem superiores a $|4,0|$, sugerem-se um grau de erro potencialmente inaceitável do modelo.

Como último passo, obtêm-se os escores fatoriais para cada observação, que possuem amplas formas de análises. São obtidas através do produto da matriz de cargas fatoriais com a matriz de dados normalizados originais, segundo Perobelli et al. (1999). Por conseguinte, os escores são as realizações das variáveis originais no modelo selecionado. Também podem ser utilizados para classificar as observações no que concernem às interpretações práticas de cada fator. Na classificação dos municípios este será um dos passos mais importantes deste trabalho.

Todo o procedimento estatístico descrito anteriormente acerca da análise fatorial foi efetuado por meio do *software* R (R Development Core Team, 2013).

3 Resultados e discussão

Em um primeiro momento, com o intuito de estudar exploratoriamente a estrutura de associação entre as 8 variáveis adimensionais mensuradas nos 30 municípios do Estado de Minas Gerais e realizar a padronização dos dados, foi calculada a matriz de correlações apresentada na Tabela 1.

Tabela 1 - Matriz de correlações entre as 8 variáveis referentes aos 30 municípios de Minas Gerais

Variável	IVAA	IVAI	IVAS	IPIB	ICFI	IGINI	IAF	IEF
IVAA	1,000							
IVAI	0,065	1,000						
IVAS	0,075	0,630	1,000					
IPIB	0,255	0,825	0,833	1,000				
ICFI	-0,182	0,047	0,160	0,105	1,000			
IGINI	0,161	-0,061	-0,179	-0,113	-0,992	1,000		
IAF	-0,242	0,019	0,179	0,184	-0,004	0,017	1,000	
IEF	0,377	0,187	0,081	0,211	0,195	-0,189	-0,544	1,000

De acordo com a matriz de correlações apresentada na Tabela 1, observa-se que existem correlações fortes e positivas entre as variáveis IPIB e IVAI, IPIB e IVAS, IVAI e IVAS, além de uma correlação muito forte e negativa entre as variáveis IGINI e ICFI.

Como a maioria das variáveis apresentaram correlações entre si, pode-se pressupor inicialmente, em tese, que o uso da análise fatorial dar-se-á de forma promissora. Essas correlações também podem ser visualizadas através de um gráfico denominado de correlograma (Figura 1).

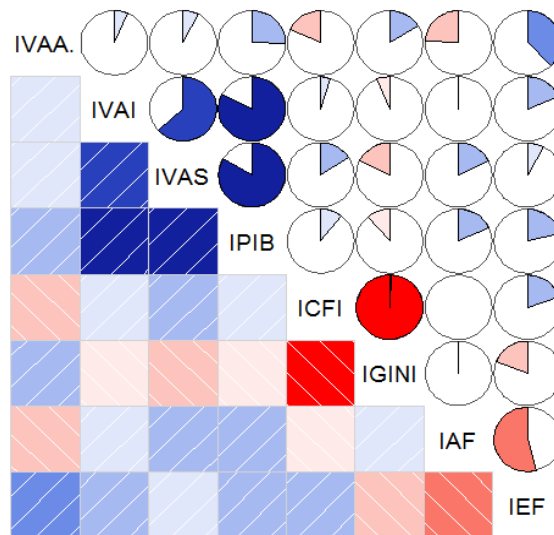


Figura 1 - Gráfico das correlações amostrais das 8 variáveis referentes aos 30 municípios do Estado de Minas Gerais.

Ainda na Figura 1, o correlograma procura descrever as correlações bivariadas de forma mais interpretativa, em que as correlações positivas estão em cor azul, com tons mais fortes para as correlações mais altas. Para estas, o ângulo (sentido horário) no gráfico de setores do painel superior é maior entre as variáveis IPIB e IVAI, IPIB e IVAS, IVAI e IVAS (correlações fortes e positivas). Já as correlações negativas estão em cor rosa, com tons mais fortes para as correlações mais altas. Para estas, o ângulo (sentido anti-horário) no gráfico de setores do painel superior é maior entre as variáveis IGINI e ICFI (correlação muito forte e negativa).

A estrutura de correlação dos dados parece indicar que o uso análise fatorial será adequado. Porém, para confirmar a consistência dos dados para a análise fatorial será usado o teste de esfericidade de Bartlett (Tabela 2).

Tabela 2 - Resultados do teste de esfericidade de Bartlett considerando as correlações bivariadas entre as 8 variáveis estudadas

Aprox. qui-quadrado	204
Grau de liberdade	28
Valor-p	<0,0001

Pela Tabela 2, vê-se pela estatística Qui-quadrado do teste de Bartlett que a hipótese nula foi rejeitada, ou seja, a matriz de correlações é estatisticamente diferente da matriz identidade ao nível de 1% de probabilidade. Devido ao baixo valor obtido para o valor-p torna-se improvável que a matriz de correlações possua a forma da matriz identidade. Desta forma, pode-se pressupor que o uso da análise fatorial torna-se pertinente, pois as correlações entre as variáveis foram significativas.

Prosseguindo com a análise fatorial, realiza-se a análise com o número de fatores igual ao número de variáveis originais (oito variáveis), de tal modo a verificar qual o percentual da variabilidade dos dados é explicada por cada fator.

Na Tabela 3, usando o método dos componentes principais, foram apresentados os autovalores, o percentual de explicação pertinente a cada fator em relação à variabilidade total e o percentual de variância acumulada.

Tabela 3 - Autovalores, percentual de variância e percentual de variância acumulada para os oito fatores

Fator	Autovalor	Perc. de variância	Perc. de variância acumulada
1	2,350	0,294	0,294
2	2,232	0,279	0,573
3	1,383	0,173	0,746
4	0,704	0,088	0,834
5	0,534	0,067	0,901
6	0,363	0,046	0,947
7	0,273	0,034	0,981
8	0,154	0,019	1

Pela Tabela 3, conforme o percentual de variância acumulada, observa-se que devem ser usados 3 fatores, pois eles contemplam mais de 70% da variabilidade total, sendo um raciocínio análogo utilizado pelos autores Perobelli et al. (1999) ao também estudar o índice de desenvolvimento econômico de alguns municípios da Zona da Mata Mineira.

Com o intuito de verificar o quanto da variabilidade das variáveis originais é explicado pelos fatores comuns, estão apresentadas na Tabela 4 as comunalidades.

Tabela 4 - Comunalidades para as 8 variáveis analisadas

Variável	Comunalidades
IVAA	0,701
IVAI	0,683
IVAS	0,710
IPIB	0,995
ICFI	0,993
IGINI	0,987
IAF	0,493
IEF	0,806

Observa-se pela Tabela 4 que todas as comunalidades foram altas, acima de 70%, exceto as variáveis IVAI e IAF, o que evidencia que mais de 70% da variabilidade das

respectivas variáveis originais são explicadas pelos 3 fatores comuns. Resultado que também endossa a adequabilidade da análise fatorial perante aos dados analisados.

No intuito de melhorar a interpretabilidade dos fatores, a análise fatorial também foi elaborada utilizando-se as rotações Varimax e Promax. A Tabela 5 apresenta os autovalores, a proporção da variância explicada e a proporção da variância acumulada explicada pelos 3 fatores ajustados sem rotação e com rotação (Varimax e Promax).

Tabela 5 - Autovalores, proporção da variância explicada e proporção da variância acumulada explicada pelos 3 fatores considerando a análise fatorial sem rotação, com rotação Varimax e com rotação Promax

	Sem rotação			Varimax			Promax		
	Fator1	Fator2	Fator3	Fator1	Fator2	Fator3	Fator1	Fator2	Fator3
Autovalores	2,350	2,232	1,383	2,502	2,059	1,403	2,494	2,049	1,413
Proporção da variância	0,294	0,279	0,173	0,313	0,257	0,175	0,312	0,256	0,177
Proporção da variância Acumulada	0,294	0,573	0,746	0,313	0,570	0,746	0,312	0,568	0,744

Pela Tabela 5, observa-se que o fator 1, sem rotação, é responsável por 29,40% (com rotação - Varimax 31,30% e Promax 31,20%) da variação total explicada pelos dados, o fator 2 sem rotação é responsável por 27,90% (com rotação - Varimax 25,70% e Promax 25,60%) e o fator 3 sem rotação é responsável por 17,30% (com rotação - Varimax 17,50% e Promax 17,70%).

A Tabela 6 apresenta as cargas fatoriais estimadas para cada fator (sem e com rotação), que são as correlações das variáveis originais com os fatores, além do critério de qualidade de ajuste Raiz Quadrada do Erro Quadrático Médio (REQM).

Tabela 6 - Cargas fatoriais estimadas do modelo fatorial ajustado para cada fator e o critério de qualidade de ajuste Raiz Quadrada do Erro Quadrático Médio (REQM)

Variável	Sem rotação			Varimax			Promax		
	Fator1	Fator2	Fator3	Fator1	Fator2	Fator3	Fator1	Fator2	Fator3
IVAI	0,795	0,229	0,037	0,823	-0,003	0,094	0,823	0,040	0,068
IVAS	0,765	0,340	-1,123	0,836	0,113	-0,064	0,837	-0,084	-0,084
IPIB	0,945	0,319	-0,003	0,994	0,042	0,067	0,994	0,000	0,038
ICFI	-0,218	0,973	-0,002	0,063	0,995	0,013	0,011	-0,995	0,069
IGINI	0,208	-0,973	0,003	-0,072	-0,992	-0,013	-0,02	0,991	-0,069
IAF	0,181	0,030	-0,676	0,229	-0,015	-0,662	0,257	-0,005	-0,671
IEF	0,147	0,233	0,854	0,146	0,174	0,868	0,103	-0,130	0,874
IVAA	0,306	-0,112	0,425	0,233	-0,198	0,440	0,226	0,227	0,421
REQM		0,224			0,224			0,346	

Os valores em negrito representam as variáveis com maior carga (em valor absoluto) dentre os fatores extraídos.

Pela Tabela 6, observa-se que os critérios de qualidade de ajuste para o modelo fatorial sem rotação e com rotação Varimax apresentaram menor REQM em relação ao critério de rotação Promax. Portanto, para facilitar a interpretação das cargas fatoriais usa-se o critério de rotação Varimax (MARDIA et al., 2006; Johnson & Wichern, 2007). Também observa-se na Tabela 6, que no fator 1 as variáveis IVAI, IVAS e IPIB estão mais correlacionadas com esse fator, as variáveis ICFI e IGINI estão correlacionadas com o fator 2 e as variáveis IAF, IEF e IVAA estão correlacionadas com o fator 3. Desse modo, o fator 1 será denominado de “indicador econômico” para os municípios, o fator 2 de “indicador fundiário” e o fator 3 de “indicador agropecuário”.

Após definição do modelo a ser utilizado, são apresentados os resíduos padronizados na Tabela 7 com o intuito de verificar a qualidade do ajuste do modelo selecionado.

Tabela 7 - Matriz de resíduos padronizados do modelo ajustado pelo critério de rotação varimax entre as 8 variáveis aferidas referentes aos 30 municípios de Minas Gerais

Variável	IVAA	IVAI	IVAS	IPIB	ICFI	IGINI	IAF	IEF
IVAA	0,000							
IVAI	-0,168	0,000						
IVAS	-0,068	-0,052	0,000					
IPIB	0,003	0,001	0,001	0,000				
ICFI	-0,006	-0,002	-0,004	0,000	0,000			
IGINI	-0,012	-0,005	-0,008	0,000	0,000	0,000		
IAF	-0,008	-0,106	-0,053	0,002	0,005	0,011	0,000	
IEF	-0,004	-0,014	-0,005	0,000	0,002	0,004	-0,001	0,000

Pela Tabela 7, percebe-se que todos os resíduos foram baixos, possuindo valores absolutos menores que 2,5. Sendo assim, o bom comportamento residual indica que não há evidências suficientes para inclusão ou exclusão de parâmetros no modelo ajustado, implicando assim em um adequado ajuste do modelo.

Tendo em vista que o fator 1 pode ser classificado como um indicador econômico, utilizam-se os escores deste fator para classificar os 30 municípios, como mostra a Figura 2. A Figura 2 mostra que, os municípios de: Ouro Branco, Urucuaia, Araxá e Extrema apresentam escores elevados, fato este que os classificam como os municípios que possuem os maiores indicadores econômicos. Por outro lado, os municípios de: São Vicente de Minas, Ipaba e Rubim apresentam escores baixos, classificando-os como municípios que apresentam menores indicadores econômicos. O fato do município de Ouro Branco possuir indicador econômico alto já era algo esperado, pois este município, segundo o PMDI (2011), está localizado na região central do estado, sendo esta muito próspera e com grande concentração de indústrias. Os demais municípios se situam em posições intermediárias nesta classificação.

Na Figura 3, utiliza-se também dos escores do fator, porém do fator 2, denominado de indicador fundiário, para classificar os 30 municípios.

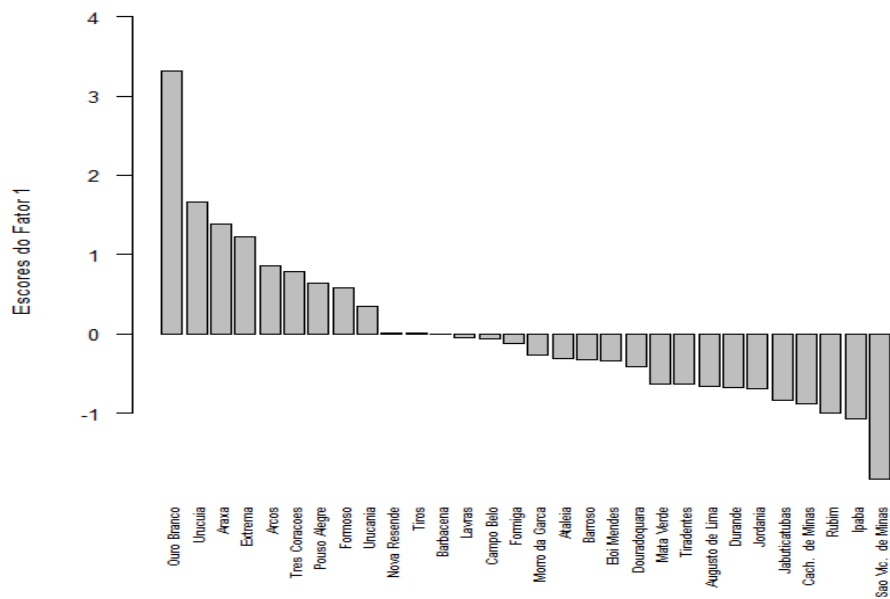


Figura 2 - Escores do primeiro fator, denominado indicador econômico, considerando a análise fatorial com rotação Varimax referente as 8 variáveis e 30 municípios.

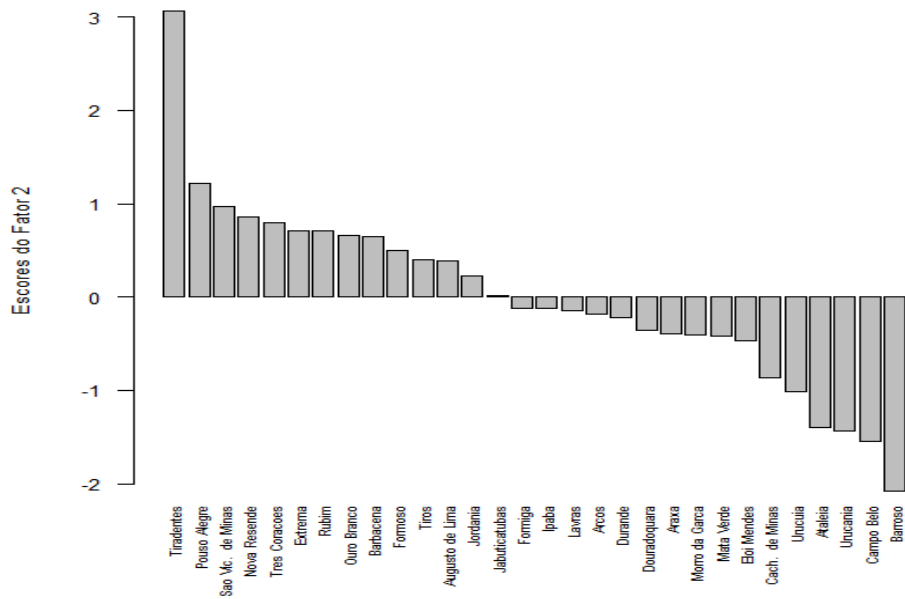


Figura 3 - Escores do segundo fator, denominado indicador fundiário, considerando a análise fatorial com rotação Varimax referente as 8 variáveis e 30 municípios.

De acordo com a Figura 3, vê-se que os municípios de: Tiradentes, Pouso Alegre e São Vicente de Minas apresentam escores elevados, fato este que os classificam como os municípios que possuem os maiores indicadores fundiários. Nesta análise quanto maior o índice, melhor distribuída está a estrutura fundiária no município. Por outro lado, os municípios de: Barroso, Campo Belo, Uruçânia e Ataléia apresentam escores baixos, classificando-os como municípios que apresentam menores indicadores fundiários, devendo assim, estas informações serem analisadas com muita ênfase, pois a má distribuição fundiária implica em consequências muito maléficas para qualquer município. O fato do município de Barroso possuir baixo indicador fundiário pode ser compreendido em função de este município possuir baixa extensão territorial, inferior a 83 km² segundo IBGE (2013).

Por fim, têm-se os escores do fator 3, denominado de indicador agropecuário, para classificar os 30 municípios, como mostra a Figura 4.

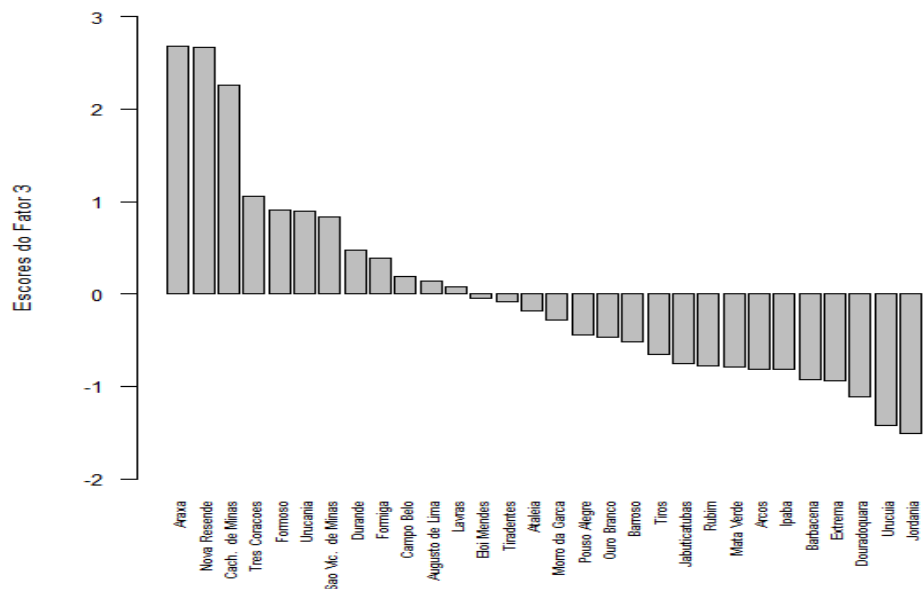


Figura 4 - Escores do terceiro fator, denominado indicador agropecuário, considerando a análise fatorial com rotação Varimax referente as 8 variáveis e 30 municípios.

Pela Figura 4, observa-se que os municípios de Araxá, Nova Resende, Cachoeira de Minas apresentam escores elevados, fato este que os classificam como os municípios que possuem os maiores indicadores agropecuários. No estudo realizado por Cunha et al. (2008), Araxá se destacou na produção de milho, por exemplo, num rol de várias cidades avaliadas no cerrado do Brasil. Por outro lado, os municípios de Jordânia, Uruçuía e Douradoquara apresentam escores baixos, classificando-os como municípios que apresentam menores indicadores de agropecuária.

Em todas as análises verificou-se que não houve destaques de municípios simultaneamente nos três indicadores, positivamente ou negativamente, entretanto, houve municípios que se enquadraram em um nível intermediário nas 3 classificações, como foi

o caso dos municípios de Lavras e Formiga, por exemplo. Observa-se também que o município de Jordânia, por exemplo, tendeu a possuir baixa classificação nos 3 indicadores.

Uma situação ideal seria que todos os municípios ocupassem posição de destaque em todos os indicadores extraídos pela análise fatorial, entretanto, vê-se que houve heterogeneidade e discrepância nas classificações dos municípios avaliados nos 3 indicadores. Consequentemente, pode-se extrair desta informação que muitas distorções ocorrem entre os 30 municípios de Minas Gerais avaliados, implicando assim que políticas públicas devem ser desenvolvidas para que estas distorções possam ser eliminadas ou, ao menos, amenizadas.

Conclusões

Pela análise das oito variáveis aferidas foi possível distinguir e caracterizar, através da análise fatorial, o perfil econômico, agropecuário e fundiário dos 30 municípios avaliados no estado de Minas Gerais.

Os resultados obtidos apontam heterogeneidade e discrepância entre os perfis caracterizados nos municípios avaliados, tornando um cenário propício para adoções de políticas públicas no intuito de apoiar maior desenvolvimento destes municípios.

Por fim, sugere-se o prosseguimento deste estudo, ampliando o número de municípios, assim como o número de variáveis aferidas, para que outros aspectos municipais/regionais possam ser captados e analisados, propiciando mecanismos e informações para que a adoção de planejamentos governamentais, ou opções empresariais, possam ser efetuadas.

SILVA, N.C.N.; FERREIRA, W.L.; CIRILLO, M.A.; SCALON, J.D. The use of factor analysis in the description and identification of the characteristic profiles of municipalities in Minas Gerais. *Rev. Bras. Biom.*, São Paulo, v.32, n.2, p.201-215, 2014.

- *ABSTRACT: Factor analysis is a multivariate statistical technique used to identify a small number of latent factors from a number of observed and correlated variables. In this work, it was used factor analysis to determine the profiles of the predominant activities in thirty municipalities of the state of Minas Gerais. For the intended study, eight variables were measured in order to subsidize and support municipal and regional planning these municipalities. Using principal axes and orthogonal rotation methods, it was possible to characterize the municipalities through three factors (economic, land and farming). The results also showed that there was heterogeneity and discrepancy among the evaluated municipalities. This scenario suggests that there is a need for actions for further development of these municipalities and, thereby, providing a favorable environment for the reversal of this situation.*
- *KEYWORDS: Multivariate analysis, economics, agriculture, factorial structure.*

Referências

BAKKE, H. A.; LEITE, A. S. M.; SILVA, L. B. Estatística multivariada: aplicação da análise fatorial na engenharia de produção. *Revista Gestão Industrial*, Ponta Grossa, v.4, n.4, p.01-14, 2008.

BROWN, T. A. *Confirmatory factor analysis for applied research*. 1ª ed., New York: The Guilford Press, 2006.

CLEMENTE, F.; GOMES, S. T. Impacto do agronegócio sobre o Índice de Desenvolvimento Sustentável (IDS) do Estado de Minas Gerais. *Revista de Política Agrícola*. Brasília, ano 20, n.4, p.69-83, 2011.

CUNHA, N. R. S.; LIMA, J. E.; GOMES, M. F. M.; BRAGA, M. J. A intensidade da exploração agropecuária como indicador da degradação ambiental na região dos Cerrados, Brasil. *Revista de Economia e Sociologia Rural*, Brasília, v.46, n.2, p.291-323, 2008.

HAIR, J. F. Jr.; ANDERSON, R. E.; TATHAM, R. L. e BLACK, W. C. *Multivariate Data Analysis with Readings*, 6ª ed., Prentice Hall, Englewood Cliffs, 2006.

Instituto Brasileiro de Geografia e Estatística – IBGE. *Contas Regionais*. Disponível em: <www.ibge.gov.br>. Acesso em: Set. 2013.

JOHNSON, R. A.; WICHERN, D. W. *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall, 2007. 773p.

MACCIOTTA, N. P. P.; CECCHINATO, A.; MELE, M.; BITTANTE, G. Use of multivariate factor analysis to define new indicator variables for milk composition and coagulation properties in Brown Swiss cows. *Journal of Dairy Science*, v.95, n.12, p.7346–7354, 2012.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. *Multivariate Analysis*. Califórnia: Academic Press, 2006. 518p.

MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG, 2005. 295p.

MOORI, R. G.; ZILBER, M. R. Um estudo da cadeia de valores com a utilização da análise fatorial. *Revista Administração contemporânea*, Curitiba, v.7, n.3, p.127-147, 2003.

NASCIMENTO, A. C. C.; LIMA, E. J.; BRAGA, M. J. et al. Eficiência técnica da atividade leiteira em Minas Gerais: uma aplicação de regressão quantílica. *Revista Brasileira de Zootecnia*. Viçosa, v.41, n.3, p.783-789, 2012.

PEROBELLI, F. S.; OLIVEIRA, A. F.; NOVY, L. G. G. et al. Planejamento regional e potenciais de desenvolvimento dos municípios de Minas Gerais na região em torno de Juiz de Fora: uma aplicação da análise fatorial. *Nova Economia*, Belo Horizonte, v.9, n.1, p.121-150, 1999.

PMDI. *Plano Mineiro de Desenvolvimento Integrado*. Belo Horizonte, 2011.

R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. www.r-project.org., 2013.

SCOLFORO, J. R. S.; OLIVEIRA, A. D.; CARVALHO, L. M. T. *Zoneamento ecológico-econômico do Estado de Minas Gerais: componentes geofísico e biótico*. Lavras: UFLA, 2008a. 161p.

SCOLFORO, J. R. S.; OLIVEIRA, A. D.; CARVALHO, L. M. T. *Zoneamento ecológico-econômico do Estado de Minas Gerais: componente socioeconômico*. Lavras: UFLA, 2008b. 195p.

SCOLFORO, J. R. S.; OLIVEIRA, A. D.; CARVALHO, L. M. T. *Zoneamento ecológico-econômico do Estado de Minas Gerais: zoneamento e cenários exploratórios*. Lavras: UFLA, 2008c. 136 p.

SEDE. *Secretaria de Estado de Desenvolvimento Econômico de Minas Gerais*. 2013. Disponível em: <<http://www.mg.gov.br/governomg/portal/m/governomg/invista-em-minas/invista-em-minas/11992-industria/11972/5042>>. Acesso em: Set. 2013.

ULLMAN, J. B. Structural Equation Modeling: *Reviewing the Basics and Moving Forward*. *Journal of Personality Assessment*, v.87, n.1, p.35-50, 2006.

Recebido em 20.11.2013

Aprovado após revisão em 08.07.2014