



SOBER

XLVI Congresso da Sociedade Brasileira de Economia,
Administração e Sociologia Rural



**UMA ABORDAGEM APLICADA DE COMPONENTES PRINCIPAIS NO ESTUDO
DO DESENVOLVIMENTO DA MESORREGIÃO METROPOLITANA DE BELO
HORIZONTE, MG**

VANESSA MARZANO ARAUJO; VANIA CORREA MOTA; IVIS BENTO DE LIMA;

UFSJ

SÃO JOÃO DEL-REI - MG - BRASIL

vanessamarzano@yahoo.com.br

PÔSTER

Desenvolvimento Rural, Territorial e regional

**UMA ABORDAGEM APLICADA DE COMPONENTES PRINCIPAIS
NO ESTUDO DO DESENVOLVIMENTO DA MESORREGIÃO
METROPOLITANA DE BELO HORIZONTE, MG**

RESUMO

Neste trabalho, buscou-se utilizar a técnica multivariada de análise dos componentes principais para verificar, dentre três indicadores selecionados, qual apresenta maior influência no nível de desenvolvimento da Mesorregião Metropolitana de Belo Horizonte para o ano de 2004. A discussão em torno do tema desenvolvimento é antiga, várias concepções já foram adotadas por teóricos, porém ainda não se conseguiu chegar a uma definição que seja clara sobre o que realmente significa desenvolvimento. A visão contemporânea do termo é que o desenvolvimento deve ir além do estritamente econômico, proporcionando o crescimento das variáveis macroeconômicas e a redistribuição dessas riquezas para toda a população, num processo que também eleve os indicadores sociais e que seja contínuo ao longo do tempo. A análise multivariada permitiu inferir que é grande a influência que a variável renda exerceu, no nível de desenvolvimento da mesorregião estudada.

PALAVRAS-CHAVE

Desenvolvimento Regional, Análise Multivariada, Belo Horizonte

ABSTRACT

In this work, the multivariate technique of analysis of the main components was used to verify, among the three indicators selected, which one shows greater influence upon the development level of the metropolitan Medium region of Belo Horizonte (city) in 2004. The discussion on the development subject is not recent, several ideas have already been adopted by theorists. However, up to the present there isn't a substantial definition considering what development really means. The contemporary vision of the term is that development must go beyond strictly economic, providing the macroeconomic variables' growth, such as wealth redistribution for the entire population, at a process that also raises social indicators and is continuous over time. Multivariate analysis allowed inferring that the variable "income" exercised great influence, at the development level of the region studied.

KEY-WORDS

Regional Development, Multivariate Analysis, Belo Horizonte

1 - INTRODUÇÃO

A discussão sobre o tema desenvolvimento mostra-se antiga, ampla e complexa. Durante o século XIX ainda não se utilizava o termo desenvolvimento. Os termos utilizados eram progresso e civilização. Acreditava-se que a América Latina poderia imitar o progresso alcançado pelos Estados Unidos, particularmente, esse otimismo era representado por um grupo de intelectuais e engenheiros positivistas, adeptos de Augusto Comte e do lema "ordem e progresso", em que o progresso depende da exploração dos recursos naturais e da produção de bens.

A ideologia desenvolvimentista surge, na América Latina, em 1945. Neste contexto, o Estado deveria orientar o desenvolvimento econômico, partindo do princípio de substituição das importações, orientada para os setores de base. O conceito de subdesenvolvimento emergiu em janeiro de 1949 quando foi anunciada a luta contra a pobreza, com o objetivo de uma "estabilização do mundo sob a égide dos Estados Unidos"(Hermet, 2002:33). Pretendiam que a América Latina alcançasse o mesmo nível de desenvolvimento das sociedades mais avançadas (Primeiro Mundo) "em função da riqueza, da educação ou da influência política" (Hermet, 2002:34). Assim, os países com esta ambição, mantiveram este rumo de 1945 a 1975, e tanto os regimes democráticos quanto os militares, deixaram de levar em conta a participação efetiva das populações no processo de desenvolvimento.

Em 1960 essa teoria desenvolvimentista passou a ser questionada por Celso Furtado, e logo depois, esse modelo baseado na exportação de matérias-primas e em substituição das importações se esgotou. Para Hermet (2002:37), o principal defeito das estratégias latino americanas de desenvolvimento foi o apoio do Estado "a uma industrialização onímoda e hipertrofiada, baseada em um setor público deficitário ou em um setor privado artificialmente protegido". Outro defeito se refere as populações trabalhadoras não somarem esforços ao processo. Contudo, a lógica do Estado indutor do desenvolvimento não trouxe os resultados esperados, criando a dívida externa e trazendo retrocesso econômico na década de 1980.

Neste contexto, surge uma demanda por novos critérios de avaliação do desenvolvimento que fosse além do estritamente econômico. Em 1990, o Programa das Nações Unidas para o Desenvolvimento – PNUD utilizou o Índice de Desenvolvimento Humano – IDH, criado por Mahbub ul Haq. Este novo índice contemplou as variáveis: renda

per capita, expectativa de vida e nível de educação. De acordo com Hermet (2002), o desenvolvimento deve ser duradouro e equitativo e os indicadores econômicos devem ser caracterizados como um meio de alcançar o desenvolvimento e não um fim.

Portanto, a proposta apresentada é de um trabalho teórico-empírico, no qual utilizou-se a técnica multivariada de análise de componentes principais. Esta técnica permite verificar qual é a variável econômica que exerce maior influência no nível de desenvolvimento da mesorregião metropolitana de Belo Horizonte, dentro dos indicadores selecionados. Tais indicadores foram: X_1 = Consumo residencial per capita de energia elétrica (KWH), X_2 = Taxa escolar de atendimento de crianças e adolescentes de 7 a 17 anos (%), X_3 = Índice mineiro de responsabilidade social (IMRS), X_4 = Empregos do setor formal em relação a população de 16 a 64 anos (%) e X_5 = Percentual da população atendida por programas de saúde da família (%).

2 - FUNDAMENTAÇÃO TEÓRICA

Nesta parte do trabalho, é estabelecida uma análise do termo desenvolvimento. Assim, é apresentada uma distinção entre desenvolvimento e crescimento. Em seguida parte-se para uma análise dos conceitos de desenvolvimento regional. Por conseguinte será apresentado o Estado de Minas Gerais e a Mesorregião Metropolitana de Belo Horizonte.

2.1 – Desenvolvimento

O termo desenvolvimento apresenta um conceito amplo e pode ser entendido de várias formas, visto que os pesquisadores ainda não conseguiram obter uma definição do termo que seja considerada satisfatória. Conforme descreveu Esteve (2000), o desenvolvimento pode ser considerado o centro de uma constelação semântica incrivelmente poderosa, afinal, não existe, no pensamento moderno, nenhum outro conceito que apresenta influência comparável no modo de pensar e agir de uma sociedade.

Ainda assim, o termo desenvolvimento pode ser associado com mudança, evolução, crescimento, metamorfose e para se conseguir refletir realmente sobre este termo, Stavenhagen sugere que se façam algumas perguntas como: “desenvolvimento de onde para onde, e de quê para quê? De pequeno a grande? De atrasado a adiantado? De simples a complexo? De jovem a velho? De estático a dinâmico? De tradicional a moderno? De pobre a rico? De inferior a superior? (Stavenhagen, 1984:12)

Com relação ao uso econômico que se faz da palavra desenvolvimento, até fins dos anos de 1970, o termo era associado a uma concepção linear evolucionista e na maioria das vezes se referindo aos países que ficaram conhecidos como países de terceiro mundo, numa nomenclatura mais atual, países emergentes. Ou seja, as ex-colônias européias ainda eram atrasadas em relação ao padrão de desenvolvimento que foi imposto pelos países colonizadores. Tal atraso, normalmente, se referia à pobreza, fome, produto interno bruto baixo, entre outros, conseqüentemente essa visão de desenvolvimento era associada ao crescimento econômico. Ainda, os países do terceiro mundo eram chamados de subdesenvolvidos por não conseguirem atingir tais indicadores sócio-econômicos que os chamados países industrializados conseguiram obter.

Porém, numa concepção contemporânea de desenvolvimento, faz-se mister que haja uma diferenciação entre crescimento econômico e desenvolvimento econômico. O primeiro termo faz referência às variáveis macroeconômicas como o crescimento do produto interno bruto, renda per capita, sem levar em consideração os aspectos sociais pelos quais a população se encontra. Já o termo desenvolvimento deve ser entendido como o crescimento das variáveis

macroeconômicas, porém esse crescimento deve beneficiar a população do país, com melhores distribuições de renda, elevados indicadores sociais, redução da miséria e da fome, bons índices de educação, redução do desemprego, etc.

Amartya Sen (2003), ao analisar o processo de desenvolvimento, afirmou que ele deve ser capaz de eliminar as privações que atingem inúmeros indivíduos no mundo. Dessa forma, o autor defende as liberdades substantivas que “incluem capacidades elementares como, por exemplo, ter condições de evitar privações como a fome, a subnutrição, a morbidez evitável, e a morte prematura bem como as liberdades associadas a saber ler e fazer cálculos aritméticos, ter participações políticas e liberdades de expressão, entre outras.” (Sen, 2003:52)

2.2 – O Desenvolvimento Regional

O desenvolvimento regional pode ser descrito como um processo endógeno que é capaz de transformar as bases econômicas e instituir melhorias dos indicadores sociais locais, ou seja, elevar os níveis de qualidade de vida da população explorando as suas capacidades e potencialidades.

Geralmente, o desenvolvimento regional se apresenta ligado às iniciativas inovadoras que requerem mobilização coletiva através do uso das potencialidades locais. Assim, pode-se afirmar que uma de suas características é que uma região busca se especializar em certa atividade na qual apresenta alguma vantagem comparativa. Bernardes (2002) lembra que o estudo da teoria das vantagens comparativas teve início em fins do século XVIII e início do século XIX com Von Thünen que publicou “O Estado Isolado” e se tornou pioneiro nessas análises. Dessa forma, a teoria das vantagens comparativas se constitui em uma das várias tentativas em explicar os diferentes graus de desenvolvimento de uma região.

Neste sentido, torna-se necessário atentar para a organização espacial da economia e o motivo pelo qual determinada região se apresenta favorável para o desenvolvimento de certa atividade. Ou seja,

“Os fatores locacionais clássicos têm significativo poder de explicação das vantagens e desvantagens gerais e específicas de cada atividade socioeconômica, que as regiões têm para atrair empreendimentos (...). A distribuição das indústrias entre as regiões, particularmente as de pequeno e médio tamanhos agrícolas, industriais e comerciais, são influenciadas pela distribuição dos recursos naturais, pelas condições edafológicas, pelos custos de transporte e por outros fatores locacionais modernos que afetam as decisões da localização das indústrias de alta tecnologia” (Ferreira *apud* Bernardes, 2002:65).

As atividades econômicas, normalmente, são influenciadas por pólos. De acordo com a Teoria dos Pólos, à medida que o processo de desenvolvimento é fomentado, a tendência é que as disparidades regionais sejam reduzidas gradativamente. Ou seja, na medida em que ocorre o desenvolvimento de determinada região, proporcionado por unidades econômicas motrizes do pólo, tal região deve estar apta a irradiar o desenvolvimento para as regiões circunvizinhas, beneficiando-as. Posto isto, ganha relevância o estudo sobre o desenvolvimento sócio-econômico de Minas Gerais.

2.3 – O Estado de Minas Gerais

Indicadores econômicos demonstram que a economia mineira ocupa a terceira posição em importância no Brasil, perdendo somente para São Paulo e Rio de Janeiro, respectivamente. Simão (2004) descreve a formação econômica de Minas Gerais e afirma que ela iniciou-se a partir da exploração e do povoamento do interior brasileiro no período colonial.

Assim, o povoamento do Estado foi induzido no século XVIII pela descoberta de grandes jazidas minerais. A atividade mineradora se declinou em fins do século XVIII, cedendo espaço para a atividade agrário-exportadora, baseada no cultivo do café, que predominou até 1950. Após este período ganhou destaque a produção siderúrgica.

Minas Gerais é o Estado que apresenta o maior número de municípios, totalizando 853. Considerando a extensão territorial, em média, um município mineiro apresenta 690 Km². Todavia, são 41 municípios que possuem extensão territorial inferior a 100 Km². Já aqueles municípios com área superior a 5.000 Km² são apenas sete.

Segundo Scavazza (2003), quase a metade (45,63%) do Produto Interno Bruto - PIB mineiro em 1999 é proveniente da região Central. Já a segunda colocação pertence à região Sul, com 12,88% do PIB total do Estado. Em seguida, encontram-se a região da Mata, com 8,47% do PIB do Estado, e as regiões do Rio Doce e Triângulo, ambas com 7,98% (embora o PIB do Rio Doce, em valores absolutos, seja maior). As regiões Centro-Oeste, Norte e Alto Paranaíba participam com 4,94%, 4,73% e 3,54% do PIB estadual, respectivamente. As menores contribuições para o PIB do Estado são das regiões do Jequitinhonha/Mucuri, que é de 1,96%, e Noroeste, de 1,89%.

Com relação ao PIB por setores da economia, verifica-se, que a maior contribuição para o PIB agropecuário em 1999 é da região Sul (23,75% do PIB agropecuário do Estado). Em seguida, estão as regiões do Triângulo (14,53%), Alto Paranaíba (12,21%) e Mata (10,70%). No que tange ao PIB industrial no mesmo ano, a liderança é da região Central, com 47,07% do total do Estado. O Sul de Minas participa com 13,01% do total do PIB industrial, ficando na 2ª posição, seguido pelo Rio Doce, com 10,63%. No setor de serviços, também a região Central ocupa o 1º lugar, contribuindo com pouco mais da metade do PIB de serviços estadual (51,48%). O 2º e 3º lugares são ocupados pelo Sul (10,70%) e Mata (9,06%).

Neste sentido, é inegável que Minas Gerais tenha características sócio-econômicas desiguais. Scavazza (2003) lembra que a preocupação com tais desigualdades está expressa na Constituição Mineira, em seu artigo 41, conforme descrito abaixo.

“Art. 41 – O Estado articulará regionalmente a ação administrativa, com o objetivo de:

(...)

II - contribuir para a redução das desigualdades regionais, mediante execução articulada de planos, programas e projetos regionais e setoriais dirigidos ao desenvolvimento global das coletividades do mesmo complexo geoeconômico e social;

III – assistir os Municípios de escassas condições de propulsão socioeconômica situados na região, para que se integrem no processo de desenvolvimento.”

Portanto, para facilitar o processo da divisão territorial brasileira, o Instituto Brasileiro de Geografia e Estatística – IBGE adota uma metodologia que trabalha com o conceito de microrregiões e mesorregiões. Este órgão descreve tais conceitos da seguinte forma (2002a:16)

“As microrregiões são um conjunto de municípios, contíguos e contidos na mesma unidade da federação, agrupados com base em características do quadro natural, da organização da produção e de sua integração. As mesorregiões são um conjunto de microrregiões também agrupadas com base no quadro natural, no processo social e na rede de comunicações e de lugares”.

Sendo assim, Minas Gerais apresenta 12 mesorregiões e 66 microrregiões. As mesorregiões são: Noroeste de Minas, Norte de Minas, Jequitinhonha, Vale do Mucuri, Triangulo Mineiro/Alto do Paranaíba, Central Mineira, Metropolitana de Belo Horizonte,

Vale do Rio Doce, Oeste de Minas, Sul/Sudoeste de Minas, Campo das Vertentes e Zona da Mata. Este estudo contemplará a Mesorregião Metropolitana de Belo Horizonte.

2.4 – A Mesorregião Metropolitana de Belo Horizonte

A história da capital mineira se inicia em 1701 quando o bandeirante João Leite da Silva Ortiz fundou, na Serra de Congonhas, a Fazenda do Cerrado. Além da procura pelo ouro, também desenvolveu atividade agrícola e pecuária. Com o progresso de tal fazenda, houve a atração de mais pessoas e aos poucos foi se formando um arraial, batizado de Curral Del Rei, mais tarde, este povoado ganhou o nome de Belo Horizonte.

Com a expansão de Belo Horizonte, foi se formando a sua região metropolitana, que atualmente ocupa uma área de mais de 20 milhões de Km²; totalizando, em 2004, uma população de aproximadamente 6 milhões de pessoas. Está localizada na transição entre os Domínios da Mata Atlântica, que se estende até a parte leste da Cadeia do Espinhaço e os Domínios dos Cerrados. Apresenta um clima tropical de altitude, com inverno seco e verão chuvoso. Durante o ano, ocorrem chuvas no período de primavera, verão e outono. É considerada uma região montanhosa e uma das mais ricas do país em recursos minerais, detendo importantes reservas de ferro, ouro, manganês e calcário. Conta com grandes recursos hídricos e com inúmeras reservas naturais, entre as quais destaca-se o Parque Nacional da Serra do Cipó. Abriga também significativos núcleos históricos, como Ouro Preto, Diamantina, Serro, Tiradentes e São João Del Rei (www.consuladoporlugalmg.org.br).



Fonte: (www.consuladoporlugalmg.org.br)

De acordo com o Índice Mineiro de Responsabilidade Social, publicado pela Fundação João Pinheiro, esta mesorregião é composta por 105 municípios que dão vida à terceira região metropolitana em importância do Brasil, representando o centro político, econômico e demográfico do Estado de Minas Gerais.

3 – METODOLOGIA

Este trabalho foi elaborado utilizando-se a metodologia de análise multivariada. Mingoti (2001:22) ressalta que a análise multivariada é importante na análise de dados, pois consiste em um conjunto de métodos usados para simplificar a interpretação de grandes conjuntos de dados. Entretanto, Hair relata ainda que

“ a análise multivariada de um modo geral, refere-se a todos os métodos estatísticos que simultaneamente analisam múltiplas medidas sobre cada individuo ou objeto sob investigação. (...) Em muitos casos, técnicas multivariadas são um meio de executar em uma única análise aquilo que antes exigia múltiplas análises usando técnicas univariadas” (2005: 26).

De acordo com Sarah Evans,

“muitas vezes precisamos examinar os dados utilizando técnicas multivariadas para podermos entender plenamente a complexidade das informações a nossa disposição. Nosso maior objetivo é fazer com que os dados “falem” de uma maneira que possam ser claramente entendida e ter credibilidade”(Malhotra. 2002:570).

Uma das técnicas de análise multivariada que pode ser utilizada é o método de componentes principais. Este método, segundo Mardia *et al* (1995), pode ser definido como uma transformação ortogonal nos dados amostrados, ou seja, transforma qualquer conjunto de variáveis em outro conjunto de novas variáveis que são não relacionadas uma com as outras (1995: 214). De acordo com Johnson & Wichern

“algebricamente componentes principais nada mais é do que combinações lineares de variáveis aleatórias X_1, X_2, \dots, X_p . Geometricamente estas combinações lineares representam a seleção de um novo sistema de coordenadas obtidas por rotação do sistema original com as X_1, X_2, \dots, X_p variáveis dos eixos das coordenadas. Os novos eixos representam a direção com máxima variabilidade e fornece uma descrição mais simples e mais detalhada da estrutura de covariância”(1992: 357).

A análise de componentes principais também pode ser usada para julgar a importância das próprias variáveis originais escolhidas, ou seja, as variáveis originais com maior peso (*loadings*) na combinação linear dos primeiros componentes principais são as mais importantes do ponto de vista estatístico (Moita & Moita, 1998).

3.1 - Aplicação dos Componentes Principais

A análise de componentes principais tem sido de grande utilidade em diversas áreas como, por exemplo, economia, administração, ciências sociais, engenharia e biologia. Na literatura existem muitos trabalhos com aplicação desta metodologia, entre eles podemos citar: Moita & Moita (1998) que aplicou os componentes principais a dados químicos, a saber, composição de alimentos envolvendo teores de calorias tais como: glicídios, proteínas, lipídios, cálcio, fósforo e ferro em frutas; Silva & Padovani (2006) usou componentes principais para verificar o efeito da variabilidade de isolados de patógenos no desenvolvimento de cultivares de milho; Marques & Marques (2005) utilizou componentes principais para reduzir um número de variáveis explicativas (independentes) da economia brasileira em modelos de regressão linear múltiplas e Lopes (2001) aplicou a análise de componentes principais para determinar a confiabilidade do tempo médio de falha de um produto com tempo de vida correlacionados. Neste trabalho, a análise de componentes principais será utilizada para verificar quais são as variáveis que exercem maior influência no nível de desenvolvimento da mesorregião metropolitana de Belo Horizonte.

3.2 - Componentes Principais

Considere o vetor $X' = [X_1 \ X_2 \ \dots \ X_p]$ constituído por p componentes. Cada componente X_i , $i=1,2,3,\dots,p$ é uma variável aleatória e possuem: autovalores $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p \geq 0$, autovetores normalizados $e_1, e_2, e_3, \dots, e_p$, vetor de médias (μ) e matriz de covariância ($\Sigma_{p \times p}$) dado respectivamente por:

$$\mu = E(X) = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \quad \text{e} \quad \Sigma_{p \times p} = \text{Cov}(X_i, X_j) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix},$$

em que, $\text{Cov}(X_i, X_j) = \sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$, sendo que $i, j = 1, 2, 3, \dots, p$. No entanto quando $i=j$ a covariância corresponde a variância da variável X_i $i=1, 2, 3, \dots, p$.

Os autovalores e os autovetores devem satisfazer as seguintes pressuposições:

- i) $e_i' e_j = 0$ para todo $i \neq j$;
- ii) $e_i' e_i = 1$ para todo $i = 1, 2, 3, \dots, p$;
- iii) $\Sigma_{p \times p} e_i = \lambda_i e_i$, para todo $i = 1, 2, 3, \dots, p$;

A componente principal da i -ésima variável é representada por Y_i é definida da seguinte maneira:

$$Y_i = e_i' X = e_{i1} X_1 + e_{i2} X_2 + e_{i3} X_3 + \dots + e_{ip} X_p,$$

além disso, a covariância entre quaisquer duas componentes deve ser nula, ou seja, $\text{Cov}[Y_i, Y_j] = 0, i \neq j$ com $i, j = 1, 2, \dots, p$.

O autovalor λ_i representa a variância da componente Y_i e como $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p$, a primeira e p -ésima componente principal representa respectivamente, a maior e a menor variabilidade. Assim a proporção da variância total do vetor X que é explicada pela i -ésima componente é:

$$\frac{\text{Var}[Y_i]}{\text{VarTotal}X} = \frac{\lambda_i}{\sum_{j=1}^p \lambda_j},$$

e a correlação entre a i -ésima componente principal e a variável X_j é dada por:

$$r_{y_i, X_j} = \frac{e_{ij} \sqrt{\lambda_j}}{\sqrt{s_{jj}}},$$

em que s_{jj} é a variância amostral da variável aleatória X_j .

Em geral, as primeiras componentes principais explicam grande parte da variância total do vetor X , ou seja, a primeira componente principal detém maior informação estatística que a segunda, que por sua vez, tem mais informação que a terceira e assim por diante. Fato que beneficia o estudo das componentes principais, uma vez que restringe o número de variáveis sem perder muita informação sobre a estrutura de variâncias e covariâncias original do vetor X .

Em casos que existe uma grande discrepância entre as variâncias das variáveis X_i as componentes principais obtidas por meio da matriz de covariância são influenciadas pelas variáveis de maior variância. Nestas circunstâncias a aplicação desse procedimento pode levar a conclusões errôneas. Uma alternativa para solucionar esse problema é colocar os dados originais em uma mesma escala de medidas, isto é possível padronizando cada uma das variáveis pela sua média e desvio padrão, definida por: $Z_i = \frac{X_i - \mu_i}{\sigma_i}$, em que $\mu_i = E(X_i)$ e

$\sigma_i^2 = \text{Var}(X_i)$. Este procedimento consiste em utilizar a matriz de correlação $R_{p \times p}$ ao invés da matriz de variâncias e covariâncias $\Sigma_{p \times p}$ para obter as componentes principais. A matriz de correlação é dada pela seguinte expressão:

$$R_{p \times p} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix},$$

em que, ρ é o coeficiente de correlação definido por: $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$, $i, j = 1, 2, 3, \dots, p$. Segundo Mingoti,

“a correlação é uma medida mais adequada para avaliar o grau de relacionamento linear de duas variáveis quantitativas do que a covariância, pois seus valores estão sempre entre os valores de referência -1 e 1. Assim, quanto mais próximo de 1, mais indicação se tem que existe um relacionamento linear positivo (crescimento) entre as variáveis X_i e X_j e quanto mais próximo de -1, mais indicação se tem da existência de um relacionamento linear negativo (decréscimo). Uma correlação próxima de zero é uma indicação numérica de um não relacionamento linear entre as variáveis em questão. Além disso, é interessante observar que, ao contrário da covariância, o coeficiente de correlação é adimensional e, logo, não sofre influência das diferenças de escalas de medidas entre as variáveis (2005:30).

Para inferir sobre as matrizes $\Sigma_{p \times p}$ e $R_{p \times p}$ deve-se estimar as matrizes amostrais de covariância e correlação, respectivamente, por $S_{p \times p}$ e $r_{p \times p}$ e o vetor μ deve ser estimado pelo vetor \bar{X} . Conseqüentemente as componentes principais estimadas são obtidas utilizando os autovalores e autovetores amostrais estimados, respectivamente, $\hat{\lambda}_i$ e \hat{e}_i , $i = 1, 2, 3, \dots, p$.

Desta forma a i -ésima componente estimada \hat{Y}_i pela matriz de correlação amostral é definida pela seguinte expressão: $\hat{Y}_i = \hat{e}_{i1}Z_1 + \hat{e}_{i2}Z_2 + \hat{e}_{i3}Z_3 + \dots + \hat{e}_{ip}Z_p$.

3.3 - Análise de dados

Para estimar as variáveis foram utilizados indicadores construídos a partir de base de dados secundários publicados pela Fundação João Pinheiro (índice mineiro de responsabilidade social). O conjunto de dados é constituído por três indicadores de 105 cidades da mesorregião metropolitana de Belo Horizonte (BH) no ano de 2004. A partir destes dados pode-se estudar a característica de cada mesorregião metropolitana diante de uma análise conjunta de variáveis selecionadas.

A técnica a ser utilizada é análise de componentes principais que possibilita representar a mesorregião metropolitana de BH por índices resultantes de adequadas combinações lineares das três variáveis econômicas selecionadas. Por meio desta técnica além da informação individual de cada índice ou componentes, temos também a informação conjunta de pares de componentes mais importantes na análise. O grau de importância é obtido pela magnitude da variância explicada das três variáveis para cada componente.

As variáveis estudadas são: X_1 = Consumo residencial per capita de energia elétrica (KWH), X_2 = Taxa escolar de atendimento de crianças e adolescentes de 7 a 17 anos (%), X_3 = Índice mineiro de responsabilidade social (IMRS), X_4 = Empregos do setor formal em relação a população de 16 a 64 anos (%) e X_5 = Percentual da população atendida por programas de saúde da família (%).

4 – RESULTADOS E DISCUSSÕES

A suposição de normalidade univariada para cada variável foi verificada por meio do teste de Shapiro-Wilk. Verificou-se que as variáveis X_4 e X_5 não seguem a distribuição normal e não foram utilizadas na análise de componentes principais e as demais variáveis atenderam a suposição com êxito. Consequentemente obteve-se a matriz de variância e covariância:

$$S = \begin{bmatrix} 9188,331 & 124,504 & 4,423 \\ 124,504 & 77,678 & 0,093 \\ 4,423 & 0,093 & 0,004 \end{bmatrix}$$

Verifica-se que as variâncias (diagonal principal) das variáveis apresentam valores discrepantes. Segundo Mingoti (2005), este fato pode influenciar na classificação das cidades. Assim, a variável consumo per capita de energia elétrica exerce maior influência quanto ao nível de desenvolvimento da mesorregião metropolitana de Belo Horizonte. Tendo em vista que ela foi utilizada como *proxy* da renda, demonstrando a forte influência dessa variável no nível de desenvolvimento dessa mesorregião.

Na Tabela 1 estão apresentados os autovalores e as variâncias explicadas pelas componentes principais obtida acima.

Tabela 1: Percentual da variação total explicada pelas componentes principais

| Componente Principal | Autovalor | Explicação (%) | Explicação Acumulada (%) |
|----------------------|-----------|----------------|--------------------------|
| 1 | 9190,035 | 99,18 | 99,18 |
| 2 | 75,977 | 0,82 | 100 |
| 3 | 0,193 | 0,00 | 100 |

A variância total das componentes \hat{Y}_1 , \hat{Y}_2 e \hat{Y}_3 pode ser explicada em porcentagens dada respectivamente, 99,18%, 0,82% e 0,00%. Observa-se que as componentes \hat{Y}_1 e \hat{Y}_2 juntas representam 100% da variação total, portanto as informações mais relevantes dos dados originais estão contidas nessas duas componentes principais, enquanto que \hat{Y}_3 não contribui para variação total.

As componentes principais estimadas são:

$$\begin{aligned} \hat{Y}_1 &= 0,9999X_1 + 0,0136X_2 + 0,0005X_3 \\ \hat{Y}_2 &= 0,0136X_1 - 0,9999X_2 - 0,0004X_3 \\ \hat{Y}_3 &= -0,0005X_1 - 0,0004X_2 + 0,9999X_3. \end{aligned}$$

A matriz de correlação resultante das variáveis em estudo foi:

$$r = \begin{bmatrix} 1,000 & 0,147 & 0,722 \\ 0,147 & 1,000 & 0,165 \\ 0,722 & 0,165 & 1,000 \end{bmatrix}$$

sendo que as variáveis mais correlacionadas são consumo residencial per capita de energia elétrica (X_1) e índice mineiro de responsabilidade social (X_3) com um coeficiente de correlação de 0,722 e a variável que apresentou menor correlação foi a taxa escolar de atendimento de crianças e adolescentes de 7 a 17 anos (X_2) com relação ao consumo residencial per capita de energia elétrica (X_1) com correlação de 0,147.

As correlações resultantes entre as componentes principais e as variáveis originais estão resumidas na Tabela 2. De um modo geral o coeficiente de maior grandeza numérica da primeira componente principal é 0,99999 e refere-se a variável X_1 , que possui maior variância. O coeficiente de menor valor é igual a zero e corresponde a X_3 , que consequentemente possui a menor variância. Pode-se observar que as componentes 1 e 2 estão correlacionadas positivamente com todas variáveis em estudo, enquanto que a componente 3 está correlacionada negativamente com todas as variáveis, exceto com a X_3 .

Tabela 2: Correlações entre as componentes principais e as variáveis originais

| Variável Original | Componentes Principais | | |
|-------------------|------------------------|---------|----------|
| | 1 | 2 | 3 |
| X_1 | 0,99999 | 0,14860 | -0,71346 |
| X_2 | 0,00124 | 0,98889 | -0,06015 |
| X_3 | 0,00000 | 0,00000 | 0,68888 |

É importante observar que as componentes 2 e 3, taxa escolar de atendimento de crianças e adolescentes de 7 a 17 anos e índice mineiro de responsabilidade social são componentes de pouca importância prática.

Vale ressaltar que as componentes principais obtidas pela matriz de covariância e matriz de correlação possuem análises e interpretações similares, proporcionando assim, consistência a análise comparativa de classificações das variáveis obtidas pelos dois métodos isso ocorre quando não há muita discrepância entre as variâncias das variáveis. Os resultados das classificações cidades da mesorregião metropolitana de BH em relação as variáveis envolvidas podem ser verificadas na Tabela 3.

Observa que as dez primeiras e dez últimas cidades apresentaram algumas semelhanças quanto a classificação de acordo com os dois métodos empregados na análise e as demais cidades quando não coincidiram ficaram próximas.

Analisando os dados apresentados na Tabela 3, pode-se inferir dois resultados: com relação a matriz de covariância, as cidades de Nova Lima, Belo Horizonte, Rio Acima, Lagoa Santa, Brumadinho, Funilândia, Itabirito, Ouro Preto, Ouro Branco e Pitangui apresentaram maior nível de desenvolvimento de acordo com as variáveis selecionadas para estudo. Tendo em vista a matriz de correlação, as cidades foram: Belo Horizonte, Nova Lima, Itabira, Lagoa Santa, Itabirito, Ouro Branco, João Monlevade, Brumadinho, Rio Acima e Pará de Minas.

Ainda com relação aos dados da Tabela 3, pode-se inferir que as seis cidades que apresentam menores índices de desenvolvimento pelos dois métodos foram: Congonhas do Norte, Itaverava, Santana do Pirapama, Santo Antônio do Itambé, Serra Azul de Minas e Rio Vermelho.

Tabela 3: Classificação das cidades de acordo com análises de matriz de covariância e de correlação

| Cidades | Matriz de Covariância | | Matriz de Correlação | |
|--------------------------|------------------------|---------|------------------------|---------|
| | Escore (\hat{Y}_1) | Posição | Escore (\hat{Y}_1) | Posição |
| Alvinópolis | 239,92 | 76 | 11,15715 | 61 |
| Alvorada de Minas | 118,59 | 104** | 10,18047 | 91 |
| Araçá | 305,50 | 44 | 12,28448 | 30 |
| Baldim | 265,12 | 64 | 10,55823 | 81 |
| Barão de Cocais | 365,15 | 26 | 13,13388 | 13 |
| Bela Vista de Minas | 294,11 | 50 | 11,56242 | 47 |
| Belo Horizonte | 582,64 | 2* | 15,47196 | 1* |
| Belo Vale | 277,22 | 58 | 11,62031 | 44 |
| Betim | 309,14 | 43 | 11,62031 | 45 |
| Bom Jesus do Amparo | 276,29 | 60 | 11,70783 | 43 |
| Bonfim | 286,85 | 52 | 11,44161 | 53 |
| Brumadinho | 479,06 | 5* | 13,41204 | 8* |
| Cachoeira da Prata | 361,89 | 27 | 12,85022 | 22 |
| Caetanópolis | 340,26 | 30 | 12,03895 | 35 |
| Caeté | 378,67 | 22 | 12,20266 | 32 |
| Capim Branco | 334,05 | 31 | 10,04827 | 94 |
| Casa Grande | 197,10 | 88 | 10,548 | 82 |
| Catas Altas | 276,41 | 59 | 12,1167 | 34 |
| Catas Altas da Noruega | 184,19 | 93 | 11,27142 | 57 |
| Conceição do Mato Dentro | 201,84 | 86 | 10,13997 | 92 |
| Confins | 314,15 | 72 | 12,25434 | 31 |
| Congonhas | 386,02 | 20 | 13,174 | 12 |
| Congonhas do Norte | 142,20 | 100** | 8,321352 | 105** |
| Conselheiro Lafaiete | 410,97 | 13 | 13,18748 | 11 |
| Contagem | 390,69 | 19 | 12,75092 | 24 |
| Cordisburgo | 212,17 | 82 | 10,59696 | 80 |
| Cristiano Ottoni | 184,02 | 94 | 11,28294 | 56 |
| Crucilândia | 255,41 | 67 | 11,45822 | 52 |
| Desterro de Entre Rios | 281,59 | 54 | 10,86233 | 72 |
| Diogo de Vasconcelos | 192,82 | 91 | 9,967862 | 95 |
| Dionísio | 212,17 | 83 | 9,545485 | 100** |
| Dom Joaquim | 184,02 | 95 | 10,83104 | 73 |
| Entre Rios de Minas | 255,41 | 68 | 11,31632 | 55 |
| Esmeraldas | 281,59 | 55 | 10,324 | 87 |
| Ferros | 192,82 | 92 | 10,08774 | 93 |
| Florestal | 297,51 | 48 | 12,98124 | 19 |
| Fortuna de Minas | 413,34 | 11 | 11,88411 | 38 |
| Funilândia | 464,85 | 6* | 10,90682 | 69 |
| Ibirité | 318,95 | 38 | 9,653079 | 99** |
| Igarapé | 195,53 | 89 | 12,43116 | 26 |
| Inhaúma | 297,51 | 49 | 11,22667 | 58 |
| Itabira | 413,34 | 12 | 13,66479 | 3* |
| Itabirito | 464,85 | 7* | 13,64133 | 5* |
| Itaguara | 318,95 | 39 | 12,62302 | 25 |
| Itambé do Mato Dentro | 195,53 | 90 | 10,63703 | 77 |
| Itatiaiuçu | 315,27 | 40 | 12,42846 | 27 |
| Itaverava | 150,55 | 99** | 8,892137 | 104** |
| Jaboticatubas | 398,10 | 17 | 12,98896 | 18 |
| Jeceaba | 226,26 | 80 | 10,66662 | 75 |

| | | | | |
|-----------------------------|--------|-------|----------|-------|
| Jequitibá | 210,71 | 84 | 9,829489 | 97** |
| João Monlevade | 405,07 | 15 | 13,49117 | 7* |
| Juatuba | 366,97 | 25 | 13,01527 | 17 |
| Lagoa Santa | 481,86 | 4* | 13,65591 | 4* |
| Maravilhas | 299,51 | 45 | 10,90029 | 70 |
| Mariana | 405,11 | 14 | 13,10568 | 14 |
| Mário Campos | 278,94 | 57 | 10,65592 | 76 |
| Mateus Leme | 331,46 | 32 | 11,58285 | 46 |
| Matozinhos | 350,02 | 28 | 12,92357 | 20 |
| Moeda | 297,52 | 47 | 11,51127 | 50 |
| Morro do Pilar | 214,90 | 81 | 10,20389 | 89 |
| Nova Era | 374,94 | 24 | 13,02892 | 16 |
| Nova Lima | 620,33 | 1* | 15,05067 | 2* |
| Nova União | 279,27 | 56 | 11,09696 | 62 |
| Onça de Pitangui | 261,94 | 65 | 10,97591 | 66 |
| Ouro Branco | 427,82 | 9* | 13,55644 | 6* |
| Ouro Preto | 430,99 | 8* | 11,77326 | 41 |
| Papagaios | 320,28 | 37 | 11,77326 | 42 |
| Pará de Minas | 401,64 | 16 | 13,18779 | 10* |
| Paraopeba | 314,54 | 41 | 11,48331 | 51 |
| Passabém | 203,30 | 85 | 10,49692 | 83 |
| Pedro Leopoldo | 376,63 | 23 | 12,17807 | 33 |
| Pequi | 329,07 | 33 | 12,36906 | 28 |
| Piedade dos Gerais | 173,24 | 97** | 10,72469 | 74 |
| Pitangui | 425,31 | 10* | 13,0709 | 15 |
| Prudente de Morais | 320,32 | 36 | 12,01671 | 36 |
| Queluzita | 242,66 | 75 | 10,89367 | 71 |
| Raposos | 392,61 | 18 | 11,8456 | 39 |
| Ribeirão das Neves | 251,06 | 73 | 10,24443 | 88 |
| Rio Acima | 488,25 | 3 | 13,28605 | 9* |
| Rio Manso | 253,29 | 71 | 11,05485 | 64 |
| Rio Piracicaba | 286,20 | 53 | 11,54908 | 48 |
| Rio Vermelho | 123,15 | 103** | 9,207828 | 101** |
| Sabará | 247,85 | 74 | 10,19099 | 90 |
| Santa Bárbara | 324,85 | 35 | 12,36822 | 29 |
| Santa Luzia | 272,08 | 61 | 10,63331 | 78 |
| Santa Maria de Itabira | 253,32 | 70 | 11,54511 | 49 |
| Santana de Pirapama | 123,47 | 102** | 9,060688 | 103** |
| Santana do Riacho | 267,26 | 63 | 10,40111 | 86 |
| Santana dos Montes | 199,48 | 87 | 10,92796 | 68 |
| Santo Antonio do Itambé | 136,09 | 101** | 9,067346 | 102** |
| Santo Antônio do Rio Abaixo | 252,77 | 72 | 11,38582 | 54 |
| São Brás do Suaçuí | 344,05 | 29 | 12,81262 | 23 |
| São Domingos da Prata | 254,71 | 69 | 11,19109 | 59 |
| São Gonçalo do Rio Abaixo | 236,96 | 77 | 11,17352 | 60 |
| São Joaquim de Bicas | 287,63 | 51 | 10,9526 | 67 |
| São José da Lapa | 298,02 | 46 | 11,88738 | 37 |
| São José da Varginha | 234,57 | 78 | 11,06439 | 63 |
| São José do Goiabal | 258,55 | 66 | 10,41869 | 85 |
| São Sebastião do Rio Preto | 183,03 | 96** | 10,60932 | 79 |
| Sarzedo | 328,61 | 34 | 11,83577 | 40 |
| Serra Azul de Minas | 97,28 | 105** | 9,813697 | 98** |
| Serro | 172,79 | 98** | 9,938238 | 96** |
| Sete Lagoas | 385,53 | 21 | 12,87041 | 21 |
| Taquaraçu de Minas | 268,08 | 62 | 10,483 | 84 |
| Vespasiano | 231,51 | 79 | 11,0423 | 65 |

*classificação das melhores cidades pelos dois métodos

**classificação das piores cidades pelos dois métodos

5 – CONCLUSÃO

Neste trabalho procurou-se investigar qual é a variável econômica que exerce maior influência no nível de desenvolvimento da mesorregião metropolitana de Belo Horizonte, dentro de três indicadores selecionados. Através de uma análise estatística multivariada, utilizando-se a técnica dos componentes principais, foi possível inferir que o indicador X_1 definido como consumo residencial per capita de energia elétrica, exerceu maior influência durante o ano de 2004. Este indicador foi utilizado como *proxy* da variável renda.

O resultado alcançado está de acordo com as características da região estudada, uma vez que a mesorregião metropolitana de Belo Horizonte apresenta um complexo industrial onde se destacam os setores de mineração, siderurgia, mecânico, têxtil, entre outros. Tal região ainda é importante por apresentar uma localização geográfica privilegiada entre os grandes centros urbanos do país.

O Brasil, por ser um país capitalista, apresenta produtos e serviços como lazer, alimentos, vestuário, entre outros, que dependem diretamente da capacidade de compra do consumidor. Esses produtos são capazes de proporcionar o bem-estar e qualidade de vida para as famílias. Dessa forma, é importante que a população tenha renda suficiente para adquirir bens e serviços que não são oferecidos gratuitamente pelo Estado a fim de contribuir para a constituição de um círculo virtuoso de desenvolvimento sócio-econômico. Sendo assim, é possível afirmar que a renda apresenta uma correlação positiva com a qualidade de vida de uma população.

Todavia, cabe lembrar que o desenvolvimento não se constitui apenas na elevação dos indicadores de renda que devem ser considerados como um meio e não um fim. O crescimento das variáveis macroeconômicas deve trazer benefícios para toda a população, redistribuindo renda e também elevando os indicadores sociais, reduzindo a fome e a miséria.

6 – REFERÊNCIAS BIBLIOGRÁFICAS

BERNARDES, E. M. **Desenvolvimento do Vale do Tietê-Paraná: um enfoque de estoques de capitais**. Tese (doutorado) – Escola Superior de Agricultura Luiz de Queiroz, 2002. 152 p.

CONSULADO PORTUGAL. Informações variadas. Disponível em <<http://www.consuladportugalmg.org.br/minas//estado.htm>>. Acesso em: 18 nov. 2006.

ESTEVE, G. Desenvolvimento. In: SACHS, W. **Dicionário do desenvolvimento**. Petrópolis: Vozes, 2000.

FUNDAÇÃO JOÃO PINHEIRO. Índice Mineiro de Responsabilidade Social. Disponível em <<http://www.fjp.gov.br>>. Acesso em: 18 nov. 2006.

HAIR, Jr., ANDERSON, J.F.; TATHAM R.E.; R.L.; BLACK, W.C. **Análise Multivariada de Dados**. 5 .ed. Porto Alegre: Bookman, 2005.

HERMET, G. **Cultura e Desenvolvimento**. Vozes: Petrópolis. 2002. 204p

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo demográfico 2000**: documentação dos microdados da amostra do censo (compact disc). Rio de Janeiro, 2002a.

JOHNSON, R.A.; WICHERN, D.W. **Applied Multivariate Statistical Analysis**. Prentice-Hall International: New Jersey. 1992. 642 p.

LOPES, L.F.D. **Análise de componentes principais aplicada à confiabilidade de sistemas complexos**. Tese (Doutorado), Universidade de Santa Catarina : Florianópolis, 2001.

MALHOTRA, N.K. **Pesquisa de marketing: uma orientação aplicada**. Porto Alegre: Bookman, 2001. 719p.

MARDIA, K.V; KENT, J.T; BIBBY, J.M. **Multivariate Analysis**. Academic Press: London. 1995, 518 p.

MARQUES, J. M; MARQUES, M.A.M., **As componentes principais no descarte de variáveis em um modelo de regressão múltipla**. Revista da FAE, Curitiba, v8, n.1, 2005, p. 93-101.

MOITA NETO, J. M; MOITA, G.C. **Uma introdução à análise exploratória de dados multivariados**. Química nova, v21, n4, 1998, p. 98-113.

SCAVAZZA, Juliana Franca. **Diferenças socioeconômicas das regiões de Minas Gerais**. Setembro de 2003. 26p. Disponível em: <<http://www.almg.gov.br/bancoconhecimento/tematico/DifReg.pdf>>. Acesso em: 30 out. 2006.

SEN, Amartya. **Desenvolvimento como liberdade**. São Paulo: Companhia das Letras, 2003
SILVA, N.R; PADOVANI, C.R. **Utilização de Componentes Principais em Experimentação Agronômica**. Energia na Agricultura. Botucatu, v. 21, n. 4, 2006. p. 98-113.

SIMÃO, R. C. S. **Distribuição de renda e pobreza no Estado de Minas Gerais**. Dissertação (mestrado). Escola Superior de Agricultura Luiz de Queiroz, 2004. 112p.

STAVENHAGEN, R.. **Etnodesenvolvimento: uma dimensão ignorada no pensamento desenvolvimentista**. Rio de Janeiro: Anuário Antropológico, 1984.