

# APLICAÇÃO DE ANÁLISE ESTATÍSTICA MULTIVARIADA NA REGIONALIZAÇÃO UTILIZADA PELO PLANEJAMENTO DE TRANSPORTES: UM ESTUDO DE CASO NA AMAZÔNIA

*Maisa Sales Gama Tobias \**

**RESUMO:** este trabalho procura analisar uma região específica, caracterizada por uma grande heterogeneidade física e humana, situada na região Oeste do Pará. Propõe uma nova forma de regionalização a partir dos resultados da aplicação de análises estatísticas multivariadas, considerando a possibilidade de se incorporar os resultados ao planejamento regional de transportes, em sua etapa de zoneamento, visando captar melhor a situação social e econômica das unidades de tráfego e, posteriormente, subsidiar as etapas seguintes de análise e planejamento de investimentos, e de formulação de políticas de transportes para a região. Foi extraído da tese de doutorado da referida autora, cuja referência bibliográfica encontra-se no final deste artigo.

## 1. INTRODUÇÃO

Nas últimas décadas, o mundo vem experimentando rápidas e profundas transformações, em decorrência de avanços alcançados pela ciência e tecnologia, com a manipulação de novos materiais e a adoção de novos procedimentos técnicos levando a um novo padrão de desenvolvimento para a humanidade. Vive-se a transição da sociedade industrial para uma sociedade apoiada basicamente em informações, e tais transformações refletem-se na forma do homem planejar o seu espaço de vivência. Devido a complexidade do mundo atual, identificada pela heterogeneidade de suas regiões, o planejamento deve de alguma forma detectar essa complexidade e retratá-la, incorporando em suas análises variáveis representativas de cada aspecto considerado.

Dentro da análise regional de transportes, um procedimento básico importante, considerando a região como espaço diferenciado, é agrupar frações desse

espaço que possuam similaridades, de acordo com um conjunto de variáveis sociais, econômicas e políticas, visando obter uma simplificação de dados a serem manipulados, e uma agregação dessas frações em unidades de tráfego, sendo chamado de regionalização ou zoneamento. No entanto, esta tarefa não é fácil, pois a quantidade de variáveis é muito grande, com escalas de atuação e naturezas bem diversas.

Para tanto, há necessidade da exploração e desenvolvimento de técnicas e métodos que possam atender tal requisito, captando cada vez mais a complexidade da situação regional. Por isto, este trabalho procura analisar uma região específica, situada na região Oeste do Pará, propondo uma nova forma de regionalização a partir dos resultados da aplicação de análises estatísticas multivariadas, considerando a possibilidade de se incorporar estes resultados ao planejamento regional de transportes, em sua etapa de zoneamento, visando captar melhor a situação social, econômica e demográfica das unidades de tráfego e,

Engenheira Civil, MSc. Transportes (CCT/UFPb). Doutoranda em Engenharia de Transportes (Escola Politécnica da USP). Departamento de Engenharia Civil. Universidade da Amazônia. Belém-Pa.

posteriormente, subsidiar as etapas seguintes de análise e planejamento de investimentos, e de formulação de políticas de transportes para a região.

## 2. APRESENTAÇÃO DA ÁREA DE ESTUDO

A região Oeste do Pará, mostrada na Figura 1, historicamente é o resultado de uma colonização feita por soldados e padres portugueses, com o objetivo de garantir a posse do território amazônico. Nos transportes, essa região é caracterizada basicamente por sua potencialidade em transporte fluvial, devido a malha hidrográfica existente, e pela presença das rodovias BR 363, chamada Cuiabá-Santarém, e BR 230, chamada Transamazônica, ao longo das quais surgiram várias cidades na região.

Apesar de ser uma área de colonização muito antiga, a região Oeste do Pará apresenta baixa densidade populacional, com fortes concentrações nas regiões ribeirinhas e extensas faixas de terra pouco habitadas. Dos municípios que a compõe, Santarém é o mais desenvolvido. O restante vive de surtos esparsos de incremento da dinâmica populacional ocasionados pelos ciclos econômicos que se sucedem, ou pelos grandes projetos implantados, vez ou outra, pelo governo federal. Esse é o caso da Transamazônica, cuja construção na década de 70 deslocou milhares de pessoas de todo o Brasil para o trabalho da nova estrada e para o projeto de colonização ao qual estava

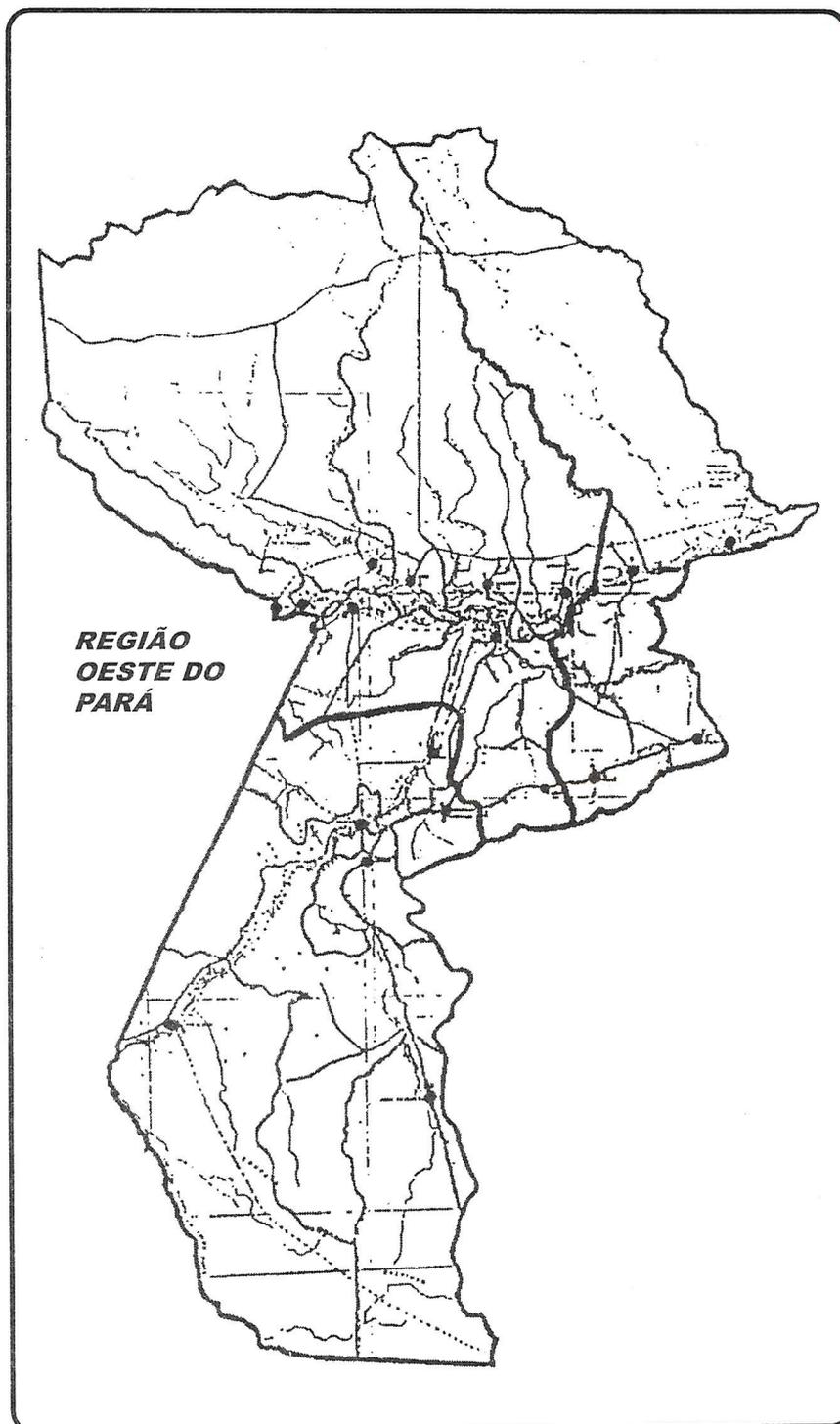


Figura 1. Área de Estudo

engajada, provocando o surgimento de novas cidades ao longo desta. Assim, essa região é constituída atualmente dos seguintes municípios: Santarém, Itaituba, Alenquer, Monte Alegre, Oriximiná, Óbidos, Almeirim, Uruará, Prainha, Juruti, Medicilândia, Rurópolis, Faro, Aveiro, Jacareacanga, Terra Santa, Trairão e Novo Progresso.

Existem algumas regionalizações da área de estudo: naturais, econômicas e ecológicas, realizadas pelo IBGE-Instituto Brasileiro de Geografia e Estatística e pelo IDESP- Instituto de Desenvolvimento Social e Econômico do Pará. A mais importante para a região Oeste do Pará foi a realizada pelo IBGE em 1968, quando implantou um novo tipo de divisão territorial em todo o Brasil, substituindo as antigas zonas fisiográficas, servindo, daí em diante, como base para a tabulação de dados estatísticos, sendo denominada de divisão por microregiões homogêneas. O estado do Pará ficou com quinze microregiões, das quais, três constituem a região Oeste do Pará, conforme é mostrado na Figura 2. Este zoneamento persiste até os dias de hoje.

Nas décadas de 70 e 80, novos incentivos aos grandes produtores agropecuários, a chegada do ciclo do ouro e a construção de grandes empreendimentos industriais e minerais, fizeram esta região estar em mutação constante e acelerada, com fortes e abruptas correntes migratórias que serviram de válvula de escape de pobreza em outras regiões, tendo recebido ao longo dessas décadas numerosos contingentes populacionais de diversos estados brasileiros. O deslocamento dessas massas migratórias em busca de frentes de trabalho acabou por ocasionar um inchamento do mercado de trabalho dos centros urbanos que não conseguiram fazer com que o número de atividades produtivas crescesse tão rapidamente quanto a sua população. Com isso, o setor de serviços, sobretudo, o de atividades não especializadas e pouco remuneradas, acabou sofrendo uma expansão rápida, ao mesmo tempo, em que esses migrantes tornaram-se invasores de terrenos e loteamentos irregulares. Essa situação tem marcado e identificado uma infra-estrutura social e econômica, de um modo geral, muito precária em todos os municípios da região Oeste do Pará.

No estudo de diversos cenários de desenvolvimento apresentados para esta região, pode-se destacar três tendências:

- ampliação da exploração mineral, organizada em grandes projetos. Já existe na região, inclusive, um grande projeto mineral, que é o projeto Trombetas, em Oriximiná;
- a perspectiva de interligação econômica com o estado do Mato Grosso, para a formação de um corredor de exportação de grãos, através da utilização da rodovia Cuiabá-Santarém e/ou implantação de uma hidrovía ligando o rio Tapajós (afluente do rio Amazonas, que atravessa a área no sentido Norte-Sul) a rios que se situam na porção Norte do Mato Grosso;
- a expansão da fronteira agrícola, que hoje se encontra no Norte do Mato Grosso e que poderá, através de melhoramentos na Cuiabá-Santarém, se expandir para a região Oeste do Pará.

Diante das modificações ocorridas nas últimas décadas, sucedendo a regionalização feita pelo IBGE, e dos cenários apresentados, torna-se relevante a realização de um estudo que vise a elaboração de uma nova proposta de zoneamento para a área de estudo, considerando as alterações nas variáveis sociais, econômicas e demográficas, para que seja analisada a situação do setor de transporte.

### 3. TÉCNICAS DE ANÁLISE MULTIVARIADA

As técnicas estatísticas multivariadas utilizadas são as Análises Fatorial e Discriminante.

#### 3.1. Considerações sobre a Análise Fatorial

A Análise Fatorial vem se tornando um termo genérico para uma variedade de procedimentos desenvolvidos com a finalidade de analisar as interrelações dentro de um conjunto de variáveis. A característica mais importante desta técnica é a sua capacidade de transformar um conjunto de dados maior em um outro

conjunto de dados menor de fatores representativos desses dados que são, em geral, não correlacionados entre si, e são extraídos do conjunto de variáveis originais, de tal forma, que o primeiro fator explique a maior quantidade de variância existente dentro do conjunto de dados, o segundo fator explique a maior quantidade de variância remanescente, isto é, a variância não explicada pelo primeiro, o terceiro fator explique parte da variância não explicada pelos dois fatores anteriores e, assim, sucessivamente. O método de extração dos fatores é chamado de método das componentes principais e faz parte do programa de Análise Fatorial da bateria de programas SPSS (Statistical Package for the Social Sciences) utilizado neste trabalho.

### 3.2. Considerações sobre a Análise Discriminante

A Análise Discriminante é uma técnica estatística de análise multivariada que tem como objetivo principal construir combinações lineares das variáveis originais que maximizem a separação dos grupos de observações. Essa técnica, também, pode ser usada para

testar se os grupos são realmente distintos, determinar o poder discriminatório das variáveis originais, identificar aquelas que mais contribuem para a separação dos grupos, avaliar a dispersão interna dos grupos, identificar elementos indevidamente classificados, alocar elementos a um dos grupos existentes e melhorar a qualidade da classificação inicial, por meio de iterações, que tem como finalidade minimizar a variância interna dos grupos e maximizar a variância entre os grupos.

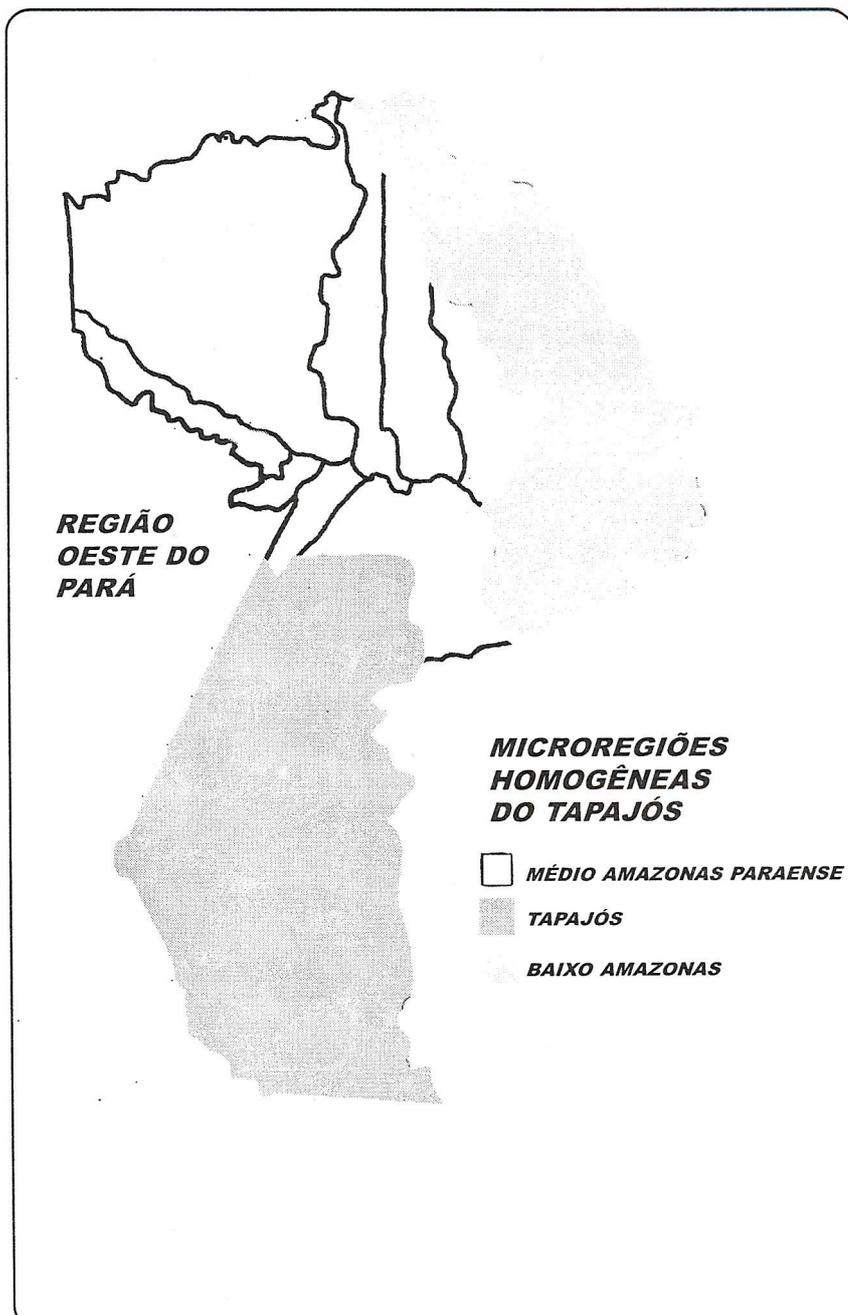


Figura 2. Zoneamento do IBGE

A técnica considera inicialmente que as "N" observações de um conjunto de "m" variáveis formam uma matriz X (N x m). Se essas N linhas podem ser arrumadas de forma a constituírem grupos de observações segundo um critério estabelecido, isto é, se há uma classificação prévia dos elementos de X, essa matriz fica dividida em matrizes de dimensões menores formada pelos grupos de observações, onde a soma do número de linhas de todas essas matrizes é igual a N. Feita essa classificação inicial, tem-se a matriz X preparada para a aplicação de uma análise discriminante. Sendo usado também o programa computacional de Análise Discriminante da bateria de programas SPSS.

#### 4. ANÁLISE E APRESENTAÇÃO DE RESULTADOS

A análise constitui-se das seguintes etapas:

**4.1. Coleta de Dados:** através do recolhimento de documentos e visitas "in loco" são selecionadas e relacionadas as seguintes variáveis para as respectivas cidades, conforme o Quadro 1.

- Vale ressaltar, que por este trabalho ter sido realizado no ano de 1995, os dados utilizados são do período de 1990 a 1995.
- A cidade de Santarém, após o ano de execução deste trabalho, já foi efetivamente desmembrada em outros dois municípios chamados de Placas (na Tranzamazônica) e Belterra (próximo a confluência do rio Amazonas com o rio Tapajós).

**4.2. Análise Fatorial:** na aplicação desta técnica são realizadas as análises fatoriais chamadas de Q e R.

**1º Passo:** é realizada uma análise fatorial do

Traços, Belém, v.2, nº 3, p.12-19, ago, 1999

#### Variáveis

X <sub>1</sub> - população	C <sub>1</sub> - Santarém
X <sub>2</sub> - domicílio	C <sub>2</sub> - Itaituba
X <sub>3</sub> - estabelecimento de saúde	C <sub>3</sub> - Alenquer
X <sub>4</sub> - empregos na saúde	C <sub>4</sub> - Monte Alegre
X <sub>5</sub> - estabelecimentos de ensino	C <sub>5</sub> - Oriximiná
X <sub>6</sub> - leitos hospitalares	C <sub>6</sub> - Óbidos
X <sub>7</sub> - empregos no ensino	C <sub>7</sub> - Almeirim
X <sub>8</sub> - vagas ocupadas na escola	C <sub>8</sub> - Uruará
X <sub>9</sub> - produção agrícola (ton)	C <sub>9</sub> - Prainha
X <sub>10</sub> - pecuária(cabeças)	C <sub>10</sub> - Juruti
X <sub>11</sub> - produção extrativa vegetal(ton/m <sup>3</sup> )	C <sub>11</sub> - Medicilândia
X <sub>12</sub> - estabelecimentos industriais	C <sub>12</sub> - Rurópolis
X <sub>13</sub> - empregos na indústria	C <sub>13</sub> - Faro
X <sub>14</sub> - estabelecimentos comerciais	C <sub>14</sub> - Aveiro
X <sub>15</sub> - empregos no comércio	C <sub>15</sub> - Jacareacanga
X <sub>16</sub> - estabelecimentos em serviços	C <sub>16</sub> - Terra Santa
X <sub>17</sub> - empregos em serviços	C <sub>17</sub> - Trairão
X <sub>18</sub> - consumo de energia elétrica residencial-rural(kwh)	C <sub>18</sub> - Novo Progresso
X <sub>19</sub> - consumo de energia elétrica comercial-industrial(kwh)	
X <sub>20</sub> - consumo de energia elétrica do setor público(kwh)	
X <sub>21</sub> - receita orçamentária dos municípios(US\$)	
X <sub>22</sub> - arrecadação estadual(R\$)	
X <sub>23</sub> - fundo de participação dos municípios(R\$)	
X <sub>24</sub> - área territorial(km <sup>2</sup> )	

Quadro 1. Coleta de Dados

tipo Q para agrupar as cidades segundo algumas dimensões econômicas importantes: i)consumo de energia elétrica residencial-rural, comercial-industrial e de utilização pelo setor público. Nesta dimensão são excluídas as cidades de Jacareacanga, Trairão e Novo Progresso, por não serem ainda energizadas pelo poder público concedente, no que já se constituem, sob este aspecto, um grupo separado das demais; ii)empregos na saúde, educação, indústria, comércio e serviços; iii)população rural e atividades produtivas: produção agrícola, extrativismo vegetal e pecuária; iv)finanças e infra-estrutura básica: estabelecimentos comerciais, industriais, de serviços, domicílios, leitos hospitalares e vaga ocupadas na escola.

**2º Passo:** executa-se uma análise fatorial do tipo Q para o agrupamento das cidades em função das 24 variáveis envolvidas, onde se pode ter uma visualização mais agregada das cidades em grupos. Analisando-se estes passos percebe-se que as cidades de Oriximiná e Almeirim apresentam tendências de agrupamento em um só e único grupo, tendo como variáveis que se sobressaem em relação às demais:  $X_4$  – empregos em saúde;  $X_7$  – empregos em ensino;  $X_{11}$  – produção extrativa vegetal;  $X_{13}$  – empregos na indústria;  $X_{17}$  – empregos em serviços e  $X_{22}$  – arrecadação estadual. Tal constatação teórica, associada ao conhecimento que se tem da realidade de cada cidade, mostra que os conjuntos de dados dos mesmos estão bastante influenciados pela presença de dois grandes projetos: o Projeto Jari, de celulose e caulim, em Almeirim e o Projeto Trombetas, de exploração de bauxita, em Oriximiná.

No entanto, observa-se que esses dois projetos em muito pouco participam da vida destas cidades, conforme consta em diversos trabalhos sobre o assunto, inclusive, estando estes projetos instalados à quilômetros de distância destas cidades, em distritos da área rural. O que leva este estudo a excluir estas cidades da análise, por este comportamento tão diferenciado e, conseqüentemente, as suas variáveis acima citadas, e a realizar-se uma análise fatorial do tipo Q para o conjunto de variáveis restantes, obtendo-se assim um agrupamento mais realista.

**3º Passo:** realiza-se uma análise fatorial do tipo R para o agrupamento das variáveis, já considerando o resultado do passo anterior, sendo que as novas variáveis se constituem da seguinte maneira:

$$X_1 = X_1 = \text{população}$$

$$X_2 = X_2 = \text{domicílio}$$

$$X_3 = X_3 + X_6 + X_{12} + X_{14} + X_{16} = \text{estabelecimentos de atividades sócio- econômicas}$$

$$X_4 = X_8 = \text{vagas ocupadas na escola}$$

$$X_5 = X_9 = \text{produção agrícola}$$

$$X_6 = X_{10} = \text{pecuária}$$

$$X_7 = X_{18} + X_{19} + X_{20} = \text{consumo de energia elétrica}$$

$$X_8 = X_{21} = \text{receita orçamentária dos municípios}$$

$$X_9 = X_5 = \text{leitões hospitalares}$$

$$X_{10} = X_{15} = \text{empregos no comércio}$$

$$X_{11} = X_{23} = \text{fundo de participação dos municípios}$$

$$X_{12} = X_{24} = \text{área territorial}$$

Uma nova Análise Fatorial é feita para agrupar as variáveis já redefinidas acima, em função das cidades, tendo-se dois fatores representativos das variáveis originais:

- $F_1 (X_1, X_2, X_3, X_4, X_5, X_6, X_9, X_{10}, X_{11}, X_7, X_8)$ , tendo como característica ser um indicador do nível de desenvolvimento urbano-rural; e
- $F_2 (X_{12})$ , tendo como característica ser um indicador de dimensão territorial.

**4º Passo:** com as variáveis do passo anterior, faz-se uma análise fatorial do tipo Q, agrupando-se as cidades em relação à essas variáveis, sendo o resultado apresentado na Figura 3.



Figura 3. Resultado da Análise Fatorial

**4.3. Análise Discriminante:** realiza-se a análise discriminante com base no agrupamento da Análise Fatorial.

O resultado desta análise mostra que as cidades de Prainha, Faro, Aveiro e Terra Santa estão indevidamente classificadas na análise anterior. Assim procede-se a Análise Discriminante, novamente, redistribuindo os referidos grupos, de acordo com as recomendações da primeira análise discriminante, tendo-se finalmente a apresentação da situação da Figura 4, com 100% de acerto nas alocações das cidades em seus respectivos grupos para as dimensões sócio-econômicas e demográficas consideradas.

Outro resultado importante da análise discriminante, é a indicação das variáveis: leitos hospitalares, empregos no comércio, estabelecimentos de atividades sócio-econômicas e pecuária como as de maior poder discriminatório entre os grupos.

## 5. CONCLUSÕES

A realização desse trabalho constata a aplicabilidade das técnicas multivariadas aos estudos de planejamento regional,

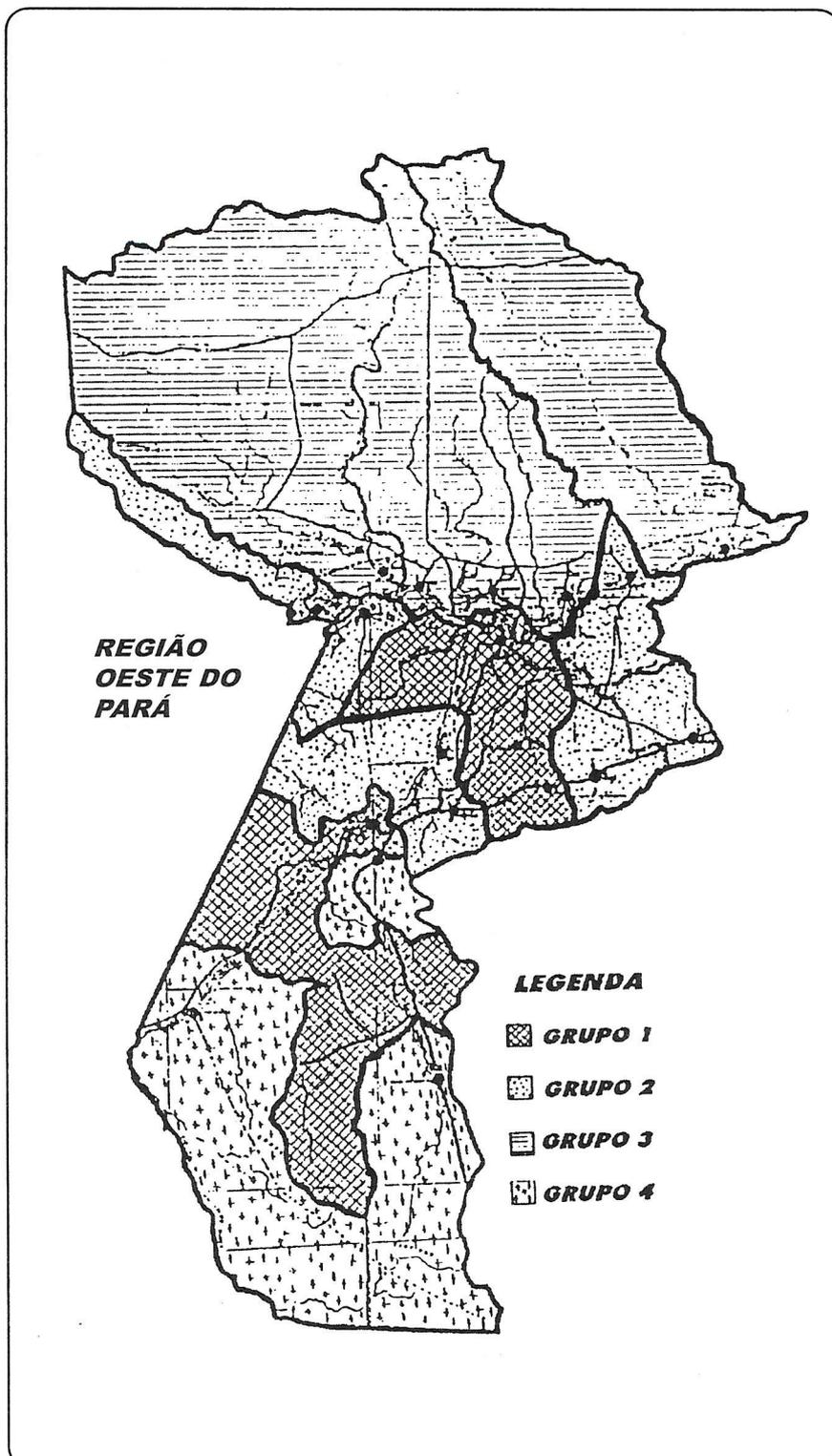


Figura 4. Resultado da Análise Discriminante

mais especificamente, na elaboração de zoneamentos com vistas ao atendimento de análises no setor de transportes e, de uma forma geral, serve para outras análises em outros setores.

O resultado final das análises Fatorial e Discriminante, que proporcionam um zoneamento em regiões homogêneas para o conjunto de variáveis sociais, econômicas e demográficas especificado, confrontado com as microregiões estabelecidas pelo IBGE, apresenta diferenças significativas no uso do solo regional, fato esse devido, principalmente, à profundas transformações experimentadas nas duas últimas décadas pela região em estudo. O resultado da Análise Fatorial, agrupando as cidades, primeiramente, segundo suas características de desenvolvimento urbano-rural e, o resultado da Análise Discriminante, mostrando que as variáveis que têm maior poder discriminatório entre as cidades referem-se às deficiências de infra-estrutura social e produtiva, levam a crer que o zoneamento encontrado neste trabalho reflete melhor a situação social, econômica e demográfica desta região, podendo servir como subsídio para estudos posteriores de planejamento regional de transportes.

## 6. BIBLIOGRAFIA CONSULTADA:

- ALMEIDA, R. S. ; RIBEIRO, M. A . C. Os sistemas de Transportes na Região Norte: Evolução e reorganização das Redes. *Revista Brasileira de Geografia*. Rio de Janeiro, v.51, nº 2, p.33-98, 1989.
- BECKER, B. K. *Amazônia*. São Paulo: Ática, 1991.
- ESTEVES, A . R. *A Ocupação da Amazônia*. São Paulo: Brasiliense, 1993.
- HARMAN, H. H. *Factor Analysis*. California: Pacific Palisades, 1968.
- MATZNETTER, J. O Sistema urbano no Norte e Nordeste do Brasil e a influência das novas Estradas. *Revista Brasileira de Geografia*. Rio de Janeiro, v.43, nº 1, p.99-122, 1981.
- PANDOLFO, C. *Amazônia Brasileira: Ocupação, desenvolvimento, e perspectivas atuais e futuras*. Belém: CEJUP, 1994.
- SUDAM, PNUD, BASA, SUFRAMA. *Macrocenários: cenários alternativos e cenários normativos da Amazônia no ano 2010*. Belém, 1990.
- SANTACRUZ, W. *Métodos estatísticos Multivariados aplicados a análise de Transportes*. Campina Grande, 1983. (Dissertação de Mestrado apresentada ao Departamento de Engenharia Civil da Universidade Federal da Paraíba, cap. 1 e 2).
- SANTOS, B. A . *Amazônia. Potencial Mineral e perspectivas de Desenvolvimento*. São Paulo: USP, 1981
- TOBIAS, M. S. *Modelo de Planejamento Integrado do Desenvolvimento Regional, Organização Espacial e Transporte para Regiões em Expansão de Fronteiras: o caso da Região Oeste do Pará*. São Paulo, 1998. (Tese de Doutorado em fase de execução no Departamento de Engenharia de Transportes da Escola Politécnica da Universidade de São Paulo).