

O NORTE DE PORTUGAL

ENSAIO DE ANÁLISE MULTIVARIADA

INTRODUÇÃO

No campo das ciências sociais os métodos de análise factorial alcançaram um amplo desenvolvimento.

Em Geografia, é crescente o recurso a estes métodos que a experiência tem mostrado serem de grande utilidade em análises a diferentes escalas do espaço — variabilidade entre países, diferenciação regional, comparação entre casos exemplares como teste de hipóteses teóricas.

Durante um estágio efectuado na Universidade de Nottingham, no ano lectivo de 1969-1970, Paula Bordalo Lema foi iniciada na análise factorial pelo Professor J. P. Cole, do Departamento de Geografia daquela Universidade, a quem expressa os seus agradecimentos. Dos ensinamentos colhidos durante o referido estágio resultou um primeiro estudo da variação das características sócio-económicas e demográficas de Portugal, em que a base espacial considerada foi o distrito (1). No trabalho que agora se apresenta, o método de análise de componentes principais é aplicado a uma escala mais restrita e utilizando informações com base no concelho. O programa para computador é da autoria de Paul M. Mather, a quem a responsável por este trabalho agradece o ter-lhe permitido utilizá-lo.

Sendo os métodos de análise factorial um conjunto de passos matemáticos, a sua aplicação implica um conhecimento

(1) P. B. LEMA e P. M. MATHER — *Factor Analysis e Cluster Analysis Aplicados a Dados Estatísticos sobre Portugal*. Nottingham University, Geography Department, 1970.

essencial do propósito desses métodos. Esta questão prévia é exposta de uma maneira breve no início deste trabalho. Os resultados obtidos nas duas experiências já realizadas — análise geral do País e análise detalhada de uma parcela do território — suscitam considerações do mesmo âmbito. Por isso, este trabalho tem o carácter de ensaio em que, para além dos resultados concretos obtidos da aplicação de um método matemático, se pretende expor criticamente as suas contribuições para a análise geográfica. Foi este o principal objectivo da sua divulgação em 1977, na série Estudos de Geografia Humana e Regional (2).

Agradecemos ao Professor Orlando Ribeiro o interesse que por ele manifestou, sugerindo-nos correcções que beneficiaram o trabalho que ora se publica.

I. O METODO DE ANALISE DE COMPONENTES PRINCIPAIS

Os métodos de análise factorial são modelos matemáticos que partem do princípio de que na variação de atributos há concomitância ou inter-relação. Concretizando, podemos considerar um exemplo simplista: ao analisar a variação, no espaço, de características económicas, sociais e demográficas, verifica-se que nas áreas mais industrializadas ocorrem um maior número de serviços de hierarquia elevada e fortes densidades de população; noutras, coincidem o fraco desenvolvimento de actividades secundárias e terciárias com baixas densidades de população.

Um dos métodos de análise factorial é o dos *componentes principais*, de que trata o presente trabalho. O seu propósito é obter, a partir do grau de associação entre variáveis, o número mínimo de componentes independentes, necessário para descrever a variação sem perda de informação. Os componentes são independentes porque cada um deles representa um conjunto diferente de variáveis com «influência» no padrão e que é expresso pela relação linear entre variáveis. Os componentes são em número mínimo porque cada um deles extrai

(2) Linha de Acção n.º 2 do Centro de Estudos Geográficos, subsidiada pelo Instituto Nacional de Investigação Científica.

o máximo da variação: o primeiro resume o padrão mais geral; da variação restante extrai-se outro factor que é, por ordem, o segundo mais geral; o poder de explicação diminui nos componentes de ordem sucessivamente mais baixa. Trata-se, pois, de reconhecer as dimensões principais da informação considerada.

A informação inicial é apresentada sob a forma de matriz em que cada linha corresponde a uma área e cada coluna a uma variável. O método de extracção dos componentes baseia-se na hipótese de que as variáveis têm distribuição normal e as relações entre elas são lineares. Nesta medida, deverão ser excluídas da matriz inicial as áreas que introduzem concentração na variação.

Obtém-se primeiramente uma matriz de correlações entre cada par de variáveis. É uma matriz simétrica em que a diagonal principal corresponde à correlação de cada variável com ela mesma (+1). A partir desta matriz, que revela o grau de inter-relação entre as variáveis, extraem-se os componentes.

Cada variável foi inicialmente padronizada de modo a corresponder a um vector com igual variância (1). A informação total é, pois, N (número de variáveis). O primeiro componente contém a máxima variação. Da variação restante extrai-se o segundo; é independente e é aquele que, logo a seguir, expressa os maiores contrastes. Obtém-se sucessivos componentes, sempre ortogonais (independentes), até que o total da variação esteja explicada ou até que a variância residual seja considerada não significativa.

O *eigen value* ou valor de cada componente indica a variação que ele contém (P). O poder de associação que ele expressa, ou seja, a percentagem de explicação que lhe corresponde é $P/N \times 100$. O valor do primeiro componente é muito elevado, decrescendo sucessivamente o dos restantes. Como cada um deles extrai sempre o máximo da variação, compreende-se que quanto maior for o grau de inter-relação das variáveis maior será o valor do primeiro componente; pelo contrário, se houver muitas variáveis pouco correlacionadas entre si, a percentagem de explicação daquele é mais reduzida e só no conjunto dos quatro ou cinco componentes se obtém uma percentagem elevada de explicação.

Os valores das variáveis da matriz inicial são ponderados e obtém-se finalmente uma nova matriz em que às colunas correspondem os componentes e, às linhas, as áreas que foram inicialmente consideradas. Esta matriz indica a posição de cada área nos componentes que são em número mais reduzido do que as variáveis iniciais — *factor scores*. Pode então interpretar-se sinteticamente a variação, ou seja, a diferenciação entre áreas.

A propriedade específica destes métodos matemáticos é reconhecer variações e regularidades. Nesta medida, compreende-se como podem constituir um importante instrumento da Geografia que tem como objecto próprio o estudo de padrões. De facto, estes modelos matemáticos das relações entre fenómenos assumem grande valor como métodos analíticos e descritivos. A sua eficácia traduz-se pela clareza e simplicidade com que permitem interpretar os factores (as influências) que um padrão geográfico complexo representa. Estas exigências não podem ser satisfeitas pela vista ou pela simples análise cartográfica. As propriedades daqueles métodos justificam as largas aplicações que têm tido em estudos geográficos. No campo da regionalização permitem reconhecer o grau de semelhança entre áreas em relação a um conjunto de variáveis; se se considerar grande número de indicadores, o objectivo é dificilmente realizável sem o recurso a estes métodos quantitativos.

As suas potencialidades no campo teórico e dedutivo são também relevantes. Abrem possibilidades amplas de experimentação, possibilitando análises múltiplas de casos exemplares ou amostras para a verificação de hipóteses e teorias. Os resultados destas experiências podem servir de base a modelos de organização espacial ou a previsões.

Neste trabalho pretende-se, com base num método de análise factorial, interpretar a variedade geográfica do Norte de Portugal, reconhecendo o grau e tipo de inter-relações que a explicam. De princípio formularam-se hipóteses acerca dos factores mais importantes de diferenciação geográfica para, com base nelas, se definir o critério de selecção dos indicadores. Efectuaram-se duas análises aplicando o mesmo método — o dos componentes principais — a um mesmo quadro espacial; elas diferem quanto aos limites territoriais e aos

indicadores utilizados. Em ambos os casos foi adoptada a base espacial do concelho. Mas na interpretação dos resultados é feita referência às divisões tradicionais de Minho e Trás-os-Montes pela relevância que elas assumem, tanto no aspecto metodológico como no prático. Ao mesmo tempo, pretendeu-se avaliar as hipóteses admitidas de início, o critério de selecção dos indicadores e o valor estatístico dos mesmos.

II. PRIMEIRA ANALISE

DISTRIBUIÇÃO DA POPULAÇÃO E LOCALIZAÇÃO DAS ACTIVIDADES ECONOMICAS

A hipótese formulada de início é a de que na variação geográfica do Norte se destacam, como factores mais importantes, a distribuição da população e o grau de desenvolvimento económico. Haveria que distinguir, além das actividades industriais, o sector de comércio de escalão mais baixo e o terciário superior. Admitiu-se ainda que a estrutura do povoamento condicionaria a densidade de estabelecimentos comerciais. Quanto à estrutura da produção agrícola, ela não é actualmente muito influenciada pela localização da indústria ou das funções urbanas mas, como a análise tinha fundamentalmente a intenção de servir de teste, considerámos que seria útil a inclusão de variáveis relativas às explorações agrícolas.

A. A INFORMAÇÃO UTILIZADA

As variáveis consideradas foram as seguintes:

1. Declive.
2. Densidade de população em 1970.
3. População presente em 1970.
4. Variação da população no decénio 1960-1970.
5. Consumo de carne por habitante em 1970.
6. Taxa bruta de natalidade em 1970.
7. Taxa de crescimento natural em 1970.
8. Taxa de emigração em 1970.
9. Assalariados nas explorações agrícolas.
10. Mecanização da agricultura.

11. Grau de comercialização e auto-consumo da produção agrícola.
12. Intensidade de utilização de fertilizantes.
13. Acessibilidade das explorações agrícolas.
14. População residente em aglomerações com mais de 1000 habitantes.
15. Número de estabelecimentos comerciais.
16. Estabelecimentos comerciais com valor de vendas superior a 100 contos.
17. Número de feiras.
18. Periodicidade máxima de realização das feiras.
19. População activa ocupada em actividades não agrícolas.
20. População activa ocupada na indústria alimentar.
21. População activa ocupada na indústria têxtil.
22. População activa ocupada em actividades terciárias — comércio, bancos e seguros.
23. População activa ocupada nos transportes.

Alguns dos índices foram elaborados por nós. Quanto à variável 1, procedemos da seguinte forma: dentro dos limites do concelho demarcámos áreas homogéneas no interior das quais se possa considerar uniforme o espaçamento das curvas de nível; para cada classe de declive assim considerada avaliámos a superfície territorial correspondente e calculámos por fim a média ponderada. Os valores das variáveis 17 e 18 foram extraídos de um estudo sobre as feiras no Norte de Portugal ⁽³⁾.

Os outros indicadores foram obtidos a partir das seguintes fontes: 2, 3, 4 e 14 — Recenseamento da população, 1970; 6, 7 e 8 — Estatísticas Demográficas, 1970; 5 — Estatísticas Agrícolas, 1970; 9, 10, 11, 12 e 13 — Inquérito às Explorações Agrícolas do Continente, 1968; 15 e 16 — Inquérito à Distribuição e Serviços, 1969; 19 a 23 — Recenseamento da população, 1960.

Como é óbvio, a forma de elaboração dos indicadores não é uniforme: 3, 17 e 18 são expressos em números absolutos;

⁽³⁾ P. B. LEMA — *As Feiras no Norte de Portugal. Contribuições para o Seu Estudo Geográfico*. Centro de Estudos Geográficos, I. A. C. Estudos de Geografia Humana e Regional, n.º 3, Lisboa 1973.

1, 4, 10, 11, 13, 14, 16, 19, 20, 21, 22 e 23 são percentagens; 2, 9 e 12 são expressos sob a forma de densidade por hectare; 5, 6, 8 e 15 são relativos a 1000 habitantes; 7 é a diferença entre a relação de nascimentos e a relação de mortos, ambos ponderados a 1000 habitantes.

Pretendeu-se analisar o Norte de Portugal limitado a sul pelo rio Douro. O concelho do Porto, bem como a área que está sob a sua influência, apresenta características sócio-económicas extremas e, por este motivo, não interessava incluir na análise. Assim, preferiu-se excluir todos os concelhos do distrito do Porto, embora reconhecendo que a influência da cidade não se estende de igual modo a toda a área do distrito.

B. ANÁLISE DE CORRELAÇÕES

Obteve-se a matriz de correlações entre todos os pares de variáveis (23×23). Por ser impossível reproduzir na íntegra a matriz de 23 variáveis, apresenta-se uma matriz reduzida que contém as 16 variáveis com grau mais elevado de correlação (fig. 1). É também uma matriz simétrica de que se reproduz uma metade e a diagonal principal. Foi elaborada por forma a que as correlações mais elevadas estivessem próximo da diagonal principal, evidenciando o grau de inter-relação das variáveis. Torna-se assim mais fácil compreender a composição dos componentes.

As variáveis relativas à emigração e às explorações agrícolas não apresentam correlações com as restantes em grau significativo. Este facto deve-se, no primeiro caso, a deficiências de informação: a fonte estatística está errada por defeito, por não incluir os contingentes clandestinos que foram muito elevados; além disso, a situação relativa a um ano não traduz as tendências de intensidade migratória. Quanto à caracterização da produção agrícola, pode dizer-se que este método não é o mais apropriado a uma regionalização agrícola, particularmente numa análise à escala do Norte de Portugal, onde se diferencia uma área com características peculiares — a região do Vinho do Porto. Poder-se-ia utilizar o indicador indirecto da produtividade por hectare que, regra geral, correlaciona com variáveis de desenvolvimento. Porém, como encobre todos os aspectos da estrutura da produção

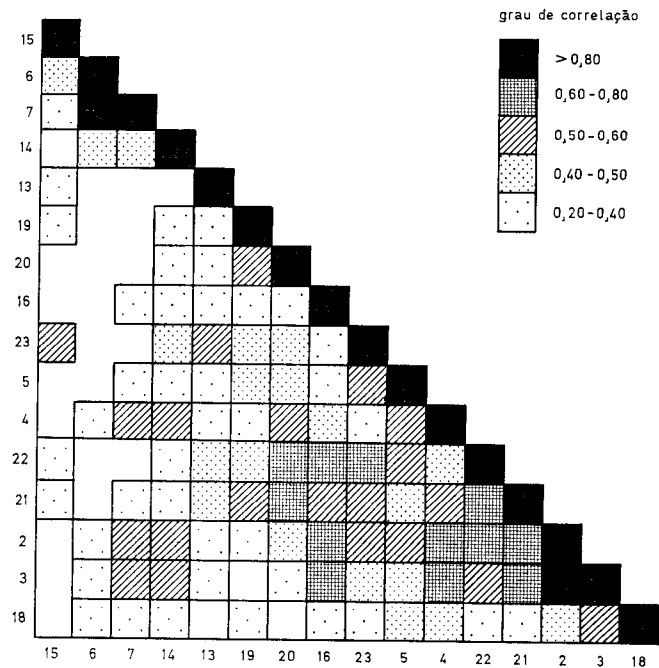
agrícola, não conduz a uma regionalização dos problemas daquela actividade. Na análise em questão não alteraria significativamente os resultados porque a área mais diferenciada é a do Vinho do Porto, como já apontámos.

Destaca-se na matriz (fig. 1) um conjunto de variáveis que apresentam correlações elevadas entre si. São as relativas à distribuição da população e às actividades da indústria e terciário (var. 2, 3, 21, 22, 23, 16, 20, 19).

Vale a pena analisar, em pormenor, as correlações entre as variáveis.

Os indicadores da distribuição da população (var. 2 e 3) estão fortemente relacionados entre si mas apresentam, com os restantes, diferentes graus de inter-relação: a densidade (var. 2) correlaciona fortemente com os transportes e o terciário (var. 23 e 22), o que não acontece com a população total (var. 3). Na realidade, as duas variáveis em questão têm significado diverso. Dois concelhos com igual população podem apresentar, devido a diferenças de área, contrastes muito grandes de densidade de população: por exemplo, o concelho de Ponte de Lima tem uma população dupla da do concelho de Montalegre mas uma densidade de população cerca de cinco vezes maior. O valor daquelas variáveis na explicação geográfica depende dos objectivos da análise.

A análise das actividades terciárias reveste-se de aspectos que interessa salientar. Os dois indicadores do equipamento comercial (var. 15 e 16) não correlacionam entre si e cada um deles tem elevada correlação apenas com uma das variáveis transportes ou terciário. Assim, a importância de estabelecimentos comerciais (var. 15) relaciona-se com o sector dos transportes e distribuição, ao passo que a variação do valor de vendas (var. 16) se explica pela maior importância do terciário superior (var. 23) e da indústria (var. 21). Na realidade, o valor de vendas é mais elevado nos concelhos de Braga e Guimarães, onde ocorre também um grande desenvolvimento de actividades secundárias. A posição imediatamente inferior na variável 16 corresponde a concelhos como Bragança e Chaves, onde a indústria não se desenvolveu, podendo concluir-se que nestes dois últimos casos a expansão de actividades comerciais, como de modo geral as do sector



Variáveis:

2. Densidade da população em 1970.
3. População presente em 1970.
4. Variação da população entre 1960 e 1970.
5. Consumo de carne por habitante em 1970.
6. Taxa bruta de natalidade em 1970.
7. Taxa de crescimento natural em 1970.
13. Acessibilidade das explorações agrícolas.
14. População residente em aglomerações com mais de 1000 habitantes.
15. Número de estabelecimentos comerciais.
16. Estabelecimentos comerciais com valor de vendas superior a 100 contos.
18. Periodicidade máxima de realização das feiras.
19. População activa ocupada em actividades não agrícolas.
20. População activa ocupada na indústria alimentar.
21. População activa ocupada na indústria têxtil.
22. População activa ocupada no comércio, bancos e seguros.
23. População activa ocupada nos transportes.

Fig. 1 — Matriz de correlações entre variáveis.

terciário, se deve ao elevado raio de atracção que aqueles centros exercem.

A variação do valor de vendas reveste-se de grande significado: reflecte uma concentração relacionada ou com a indústria ou com factores de interacção espacial. O poder de discriminação desta variável é nesta análise muito diminuta, como a seguir veremos, devido ao facto de outras variáveis incluídas apresentarem inter-relações dominantes. Haveria toda a vantagem em efectuar uma análise mais restrita, dando maior significado ao componente de grau de equipamento de comércio e serviços e utilizando indicadores mais rigorosos.

Não se evidencia relação entre comércio fixo e móvel em virtude de a periodicidade das feiras apresentar pequena variação — de uma a quatro vezes por mês. É verdade que as feiras podem realizar-se com maior frequência quando a dimensão populacional é garante de uma clientela suficiente (var. 18 e 3). Assim, no interior, nomeadamente em Trás-os-Montes, realizam-se feiras com periodicidade máxima (semanal) na sede dos concelhos onde a população é mais elevada ou as características locativas restringem a competição entre áreas de mercado. No Noroeste, com fortes densidades de população, as feiras atingiram a periodicidade máxima tanto em áreas rurais como em áreas urbanas e nestas últimas, onde as actividades de comércio fixo estão muito desenvolvidas, passaram a assumir características diferentes: na periferia da grande aglomeração do Porto, para citar um caso extremo, são essencialmente grandes mercados de produtos agrícolas. Não surpreende, pois, que a periodicidade das feiras correlacione unicamente com a população, uma vez que essa periodicidade é um carácter pouco definidor da hierarquia das feiras. Só a análise do tipo e variedade de produtos transaccionados pode revelar a relação entre esta actividade e o comércio fixo concorrente em termos de economia espacial.

Observa-se na matriz (fig. 1) que a natalidade (var. 6) não correlaciona com o conjunto de variáveis que acabámos de referir mas com o crescimento natural (var. 7) que, por sua vez, está relacionado com a distribuição da população (var. 2, 3 e 4). De facto, os contrastes sócio-económicos entre o litoral e o interior não correspondem a diferenças do mesmo grau na taxa de natalidade: esta apresenta valores muito

semelhantes em Vimioso e em concelhos do distrito de Braga. A natalidade não é, pois, um indicador sócio-económico claro e preciso. É noutros indicadores, como a estrutura etária, que se repercutem prioritariamente as diferenças de desenvolvimento económico.

As correntes migratórias estabelecem-se das áreas em estagnação para as áreas mais «ricas», sendo constituídas pelos estratos de população potencialmente activa. Este êxodo denota-se imediatamente pelo desequilíbrio na estrutura etária da população residente nas regiões de origem: fraca representação dos estratos médios (população adulta) e elevada percentagem de população idosa. Assim, e considerando ainda o exemplo acima citado de concelhos do litoral e do interior, onde as taxas de natalidade são semelhantes, pode verificar-se quanto a taxa de crescimento natural é diferenciada: os valores mais baixos em Vimioso indicam um processo de envelhecimento da população. A longo prazo são evidentes as repercussões do êxodo na base da estrutura etária. Em concelhos trasmontanos, onde o processo cumulativo de envelhecimento está mais adiantado, taxas baixas de natalidade são coincidentes com o fraco crescimento natural.

Se pensarmos que correlação significa ocorrência simultânea de dois ou mais atributos, então não surpreendem os resultados obtidos quanto às características demográficas. Eles demonstram que o desenvolvimento económico não tem influência linear e directa na natalidade. Esta conclusão não significa ausência de relações entre a evolução demográfica e as condições sócio-económicas mas evidencia, pelo contrário, como estas relações são complexas. As influências processam-se em cadeia, comportando-se a natalidade como um «produto» da interacção de outras componentes da evolução demográfica, de modo que a sua posição no processo lhe confere um tempo de evolução lento em relação aos elementos de que depende.

É neste âmbito que devem ser interpretados os resultados da aplicação de um método matemático: as correlações traduzem a situação num dado momento, enquanto o processo demográfico evolui por fases consecutivas. Num momento estatístico M_1 a natalidade não apresenta variação entre o litoral e o interior porque corresponde a um comportamento tradicional generalizado. Após um intervalo de tempo mais

ou menos longo, e durante o qual se verificaram intensas correntes migratórias das áreas mais pobres para as áreas mais ricas, regista-se em M_2 variações consideráveis no estrato médio da pirâmide de idades: aumento de população adulta nas áreas mais ricas. É natural que, decorrido outro intervalo de tempo, a observação estatística num aumento M_3 revele que outro componente apresenta variação correlacionada com o processo acima descrito: a taxa de crescimento natural, muito influenciada pela elevada percentagem de população idosa, revela-se como variável significativa, evidenciando o contraste entre áreas atractivas e repulsivas. Porém, em virtude da sua posição extrema no processo demográfico, a natalidade não mostrará ainda uma variação com igual significado. É finalmente no momento M_4 que o desequilíbrio na estrutura etária da população se vai traduzir numa descida drástica da natalidade. Pode dizer-se que esta fase do processo corresponde a um abrandamento do dinamismo dos componentes demográficos, dando lugar a um estado irreversível de estruturas demográficas bem definidas e contrastantes, que no seu conjunto se explicam pelas condições sócio-económicas.

A compreensão deste processo permite esclarecer o significado das correlações obtidas com a variável natalidade. A inclusão de outras variáveis, com a estrutura etária e os fluxos de migração interna e externa, permitiriam identificar de forma mais clara as diversas situações que definem um modelo de evolução demográfica.

C. OS COMPONENTES

QUADRO I

Valor dos primeiros 4 componentes e percentagem cumulativa de variação que contém

Componente	Valor	% Cum. variação
I	7,099	30,9
II	3,016	44,0
III	2,083	53,0
IV	1,749	60,6

QUADRO II

Peso das variáveis nos 4 componentes

1.º componente			2.º componente			3.º componente		4.º componente	
ordem	valor	variável	ordem	valor	variável	valor	variável	valor	variável
1	-0,89	2	1	-0,80	6	0,61	11	-0,68	10
2	-0,82	22	2	-0,72	7	-0,60	8	0,53	1
3	-0,82	21	3	0,69	15	-0,59	9	-0,43	8
4	-0,81	3	4	0,48	13	0,58	12	0,40	11
5	-0,77	4	5	-0,40	14	-0,38	18	0,39	23
6	-0,71	5	6	0,40	23	-0,33	16	0,37	15
7	-0,70	23	7	0,37	19	-0,20	22	0,33	14
8	-0,67	16	8	0,35	12				
9	-0,63	20	9	-0,31	4				
10	-0,58	14							
11	-0,51	18							
12	-0,51	19							
13	-0,51	13							
14	-0,50	7							

O primeiro componente resume o padrão mais geral das relações entre variáveis. A ele corresponde a máxima percentagem de explicação, como revela o quadro I. O quadro II indica que ele contém as variáveis que estabelecem os maiores contrastes — as relativas à distribuição da população (var. 2, 3 e 4) e às actividades secundárias e terciárias (var. 21 e 22). O indicador do valor de vendas é concordante com este componente; como, porém, a sua variação é mais pequena, também o seu peso no componente é muito inferior.

O segundo componente associa variáveis que se apresentavam pouco correlacionadas com aquele primeiro conjunto — natalidade e crescimento natural (var. 6 e 7), densidade de estabelecimentos comerciais (var. 15) e acessibilidade (var. 13); e ainda aquelas que são com ele concordantes mas, em virtude da elevada dispersão que apresentam, são também explicados noutra dimensão perpendicular ao primeiro componente — transportes (var. 23), grau de aglomeração (var. 14) e actividades não agrícolas (var. 19).

Atendendo às correlações das variáveis nos componentes, podemos concluir que o primeiro reflecte a oposição mais geral entre áreas urbanas e industriais do Minho, densamente

povoadas, e as áreas rurais do interior em rarefacção demográfica. O segundo explica uma menor variação, diferenciando áreas onde há elevada densidade de estabelecimentos comerciais e a natalidade da população é baixa em relação a outras onde o crescimento natural é elevado e a estrutura comercial é elementar.

O terceiro e quarto componentes contêm variáveis de emigração e da estrutura das explorações agrícolas que não correlacionam com variáveis demográficas ou actividades secundárias e terciárias, como já referimos. Elas revelam que no quadro das estruturas agrárias se destacam as áreas de agricultura não mecanizada, com base numa elevada densidade de mão-de-obra, cuja produção está orientada para o mercado, e numa intensa utilização de fertilizantes. Estas características têm o seu valor máximo na região bem diferenciada do Vinho do Porto. Os componentes traduzem este contraste mais evidente. Porém, as correlações são muito baixas, porque a inter-relação daquelas características não assume na realidade um significado fundamental, quer dizer, não define um conjunto de elementos suficiente para diferenciar de forma clara e simples a organização da actividade agrícola. Os resultados confirmam, pois, a hipótese admitida de início: as estruturas agrárias do Norte são demasiado complexas para que as variáveis consideradas possam explicar grande parte da sua variação.

A posição dos concelhos nos dois primeiros componentes foi representada num gráfico de eixos perpendiculares (fig. 2). No eixo horizontal os concelhos estão escalonados segundo o grau de crescimento demográfico e de desenvolvimento da indústria e serviços — primeiro componente. Num extremo destacam-se Braga, Guimarães, Vila Nova de Famalicão e Barcelos; seguem-se os concelhos da periferia desta área industrial e aqueles do interior onde se localizam os principais centros de comércio e serviços — Vila Real, Chaves, Bragança e Mirandela. No outro extremo estão os concelhos rurais.

O segundo componente complementa o primeiro, estabelecendo uma maior diferenciação no Minho. Assim, no eixo vertical estão os concelhos onde se verifica excesso de estabelecimentos comerciais sem relação com o desenvolvimento industrial, evolução demográfica negativa e natalidade de valor

muito baixo — são os concelhos da fronteira do Minho, onde a actividade do turismo é dominante e a emigração é intensa e até facilitada pela posição de fronteira. No outro extremo do componente estão as áreas de montanha entre o Minho e Trás-os-Montes, onde a natalidade mantém valores elevados e o equipamento comercial é muito fraco.

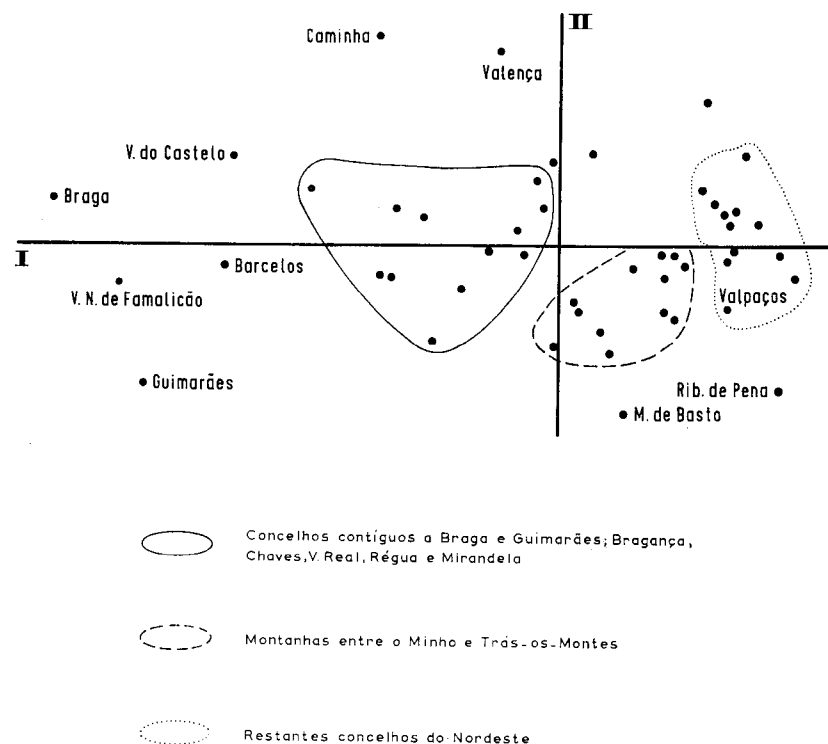


Fig. 2 — Posição dos concelhos nos dois primeiros componentes.

Do terceiro e quarto componentes não se obtém um padrão simples de interpretar, pelas razões que já expusemos. Destacam-se os concelhos durienses, particularmente Régua e Alijó, que marcam os limites da área onde a agricultura do Vinho do Porto tem maior intensidade.

D. AGRUPAMENTO DE CONCELHOS

A partir da posição coordenada dos concelhos nos componentes (ou dimensões) pode calcular-se, num espaço euclidiano, a distância entre cada par de pontos e considerá-la

Bragança
 Mirandela
 Chaves
 Monção
 Vila Real
 Fafe
 Amares
 Esposende
 Póvoa de Lanhoso
 Ponte de Lima
 Vila Verde
 Régua
 Mesão Frio
 Paredes de Coura
 Vila Nova de Cerveira
 Caminha
 Valença
 Vimioso
 Miranda do Douro
 Mogadouro
 Montalegre
 Vinhais
 Alfândega da Fé
 Murça
 Vila Flor
 Boticas
 Macedo de Cavaleiros
 Valpaços
 S.Marta de Penaguião
 Torre de Moncorvo
 Ede Espada-à-Cinta
 Carrazeda de Ansiães
 Sabrosa
 V.Pouca de Aguiar
 Celorico de Basto
 Cabeceiras de Basto
 Alijó
 Vieira do Minho
 Terras do Bouro
 Ponte da Barca
 Arcos de Valdevez
 Melgaço
 Ribeira de Pena
 Mondim de Basto
 Guimarães
 Braga
 V.N. de Famalicão
 Barcelos
 Viana do Castelo

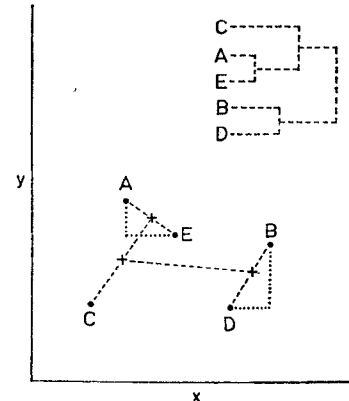
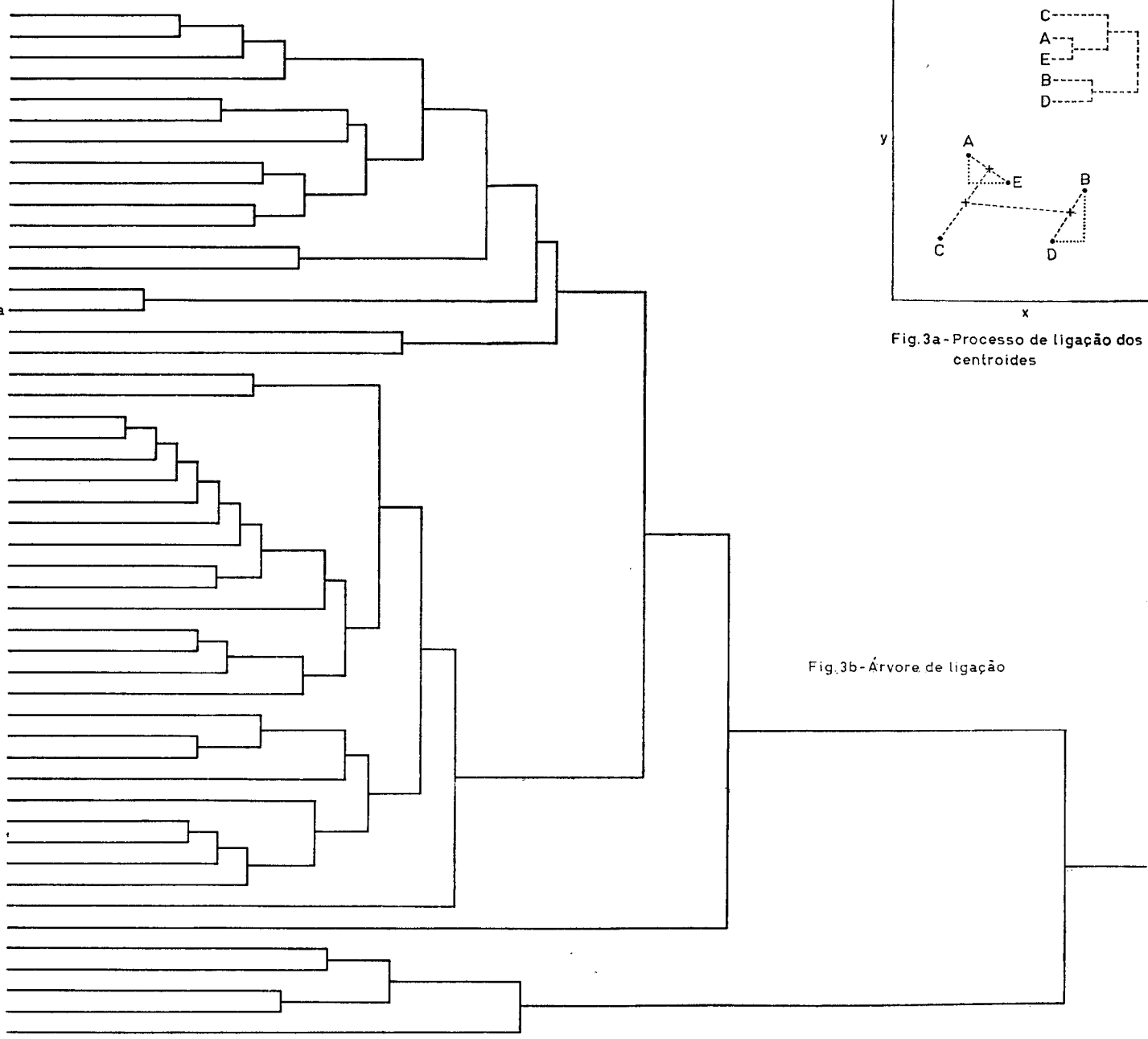


Fig.3a- Processo de ligação dos centroides

Fig.3b- Árvore de ligação

indicadora do grau de semelhança. Por este processo é possível agrupar os concelhos em áreas ou regiões homogéneas ⁽⁴⁾. Os pontos mais próximos correspondem a concelhos que apresentam maior grau de semelhança e serão eles os primeiros a formar um grupo homogéneo. O centro deste é ligado ao de outros, seguindo sempre o princípio da distância mais curta (fig. 3-a). Pela ligação subsequente dos centróides obtém-se uma árvore em que o número de grupos se vai reduzindo à medida que o grau de semelhança é menor (fig. 3-b).

O agrupamento em 11 áreas homogéneas (fig. 4) é muito concordante com a posição dos concelhos nos dois primeiros componentes extraídos, uma vez que são estes que mais contribuem para explicar a variação. Diferencia-se uma área industrial limitada aos concelhos de Braga, Guimarães, Vila Nova de Famalicão e Barcelos, onde se concentram os níveis mais elevados do sector terciário, e Viana do Castelo, devido à posição hierárquica da cidade na rede urbana do Noroeste. Os concelhos contíguos apresentam maior afinidade com os concelhos trasmontanos onde se localizam os centros de comércio e serviços mais importantes — Vila Real, Chaves, Bragança e Mirandela. Os restantes concelhos trasmontanos e as montanhas do Minho são semelhantes entre si mas individualizam-se porque, embora a predominância de actividades agrícolas seja um factor comum, as diferenças no comportamento demográfico da população são muito grandes.

III. SEGUNDA ANALISE

DINAMISMO DEMOGRAFICO E CONDIÇÕES SÓCIO-ECONÓMICAS

Interessa aprofundar o problema do modelo de evolução demográfica e em que medida ele é influenciado por condicionalismos sócio-económicos. Antes de mais, torna-se necessário avaliar o significado dos indicadores demográficos, considerando o resultado de análises já efectuadas.

⁽⁴⁾ Veja-se J. P. COLE e C. M. KING — *Quantitative Geography*, cap. 6 e 7, Londres 1970.

Quanto à natalidade, a primeira análise demonstrou não estar ela relacionada com o grau de desenvolvimento económico. De uma análise multivariada aplicada a todo o País, e em que se incluíam a natalidade e variáveis sócio-económicas, obtiveram-se resultados ainda mais esclarecedores: o factor demográfico é independente do factor desenvolvimento económico; enquanto este estabelece a diferenciação entre o litoral e o interior, o primeiro descreve a oposição geral entre o Sul e o Norte, onde a natalidade atinge os valores mais elevados de continente ⁽⁵⁾.

É certo que a natalidade não corresponde de forma mais exacta ao ritmo de nascimentos, porque pondera toda a população, independentemente do seu grupo etário. Nesta medida, a taxa de fertilidade é um indicador mais preciso. A análise de aspectos da demografia portuguesa relativa à década de 1950-1960 ⁽⁶⁾ confirma que aqueles dois índices conferem resultados diferentes. No Norte as taxas de fertilidade mais baixas registavam-se nos distritos de Viana do Castelo e Guarda, enquanto Bragança e Vila Real, numa posição intermédia, se aproximavam dos valores de Braga. Porém, na taxa bruta de natalidade, a posição daqueles dois distritos trasmontanos era inferior, devido ao peso da população idosa em relação à população adulta. Ressalvando estas diferenças, a fertilidade não tem um padrão distinto do da natalidade: estabelece da mesma forma um contraste Norte-Sul.

Na taxa de mortalidade relativa à população com idade superior a 5 anos, por forma a excluir os óbitos infantis, é ainda maior a influência da estrutura etária. Não surpreende, pois, que seja mais significativa a sua correlação com o índice de envelhecimento (relação entre a população dos grupos de mais de 60 e menos de 20 anos). É a variação deste último índice que traduz de forma mais directa o dinamismo demográfico:

Braga	21	Aveiro	27
Vila Real	23	Viseu	31
Porto	24	V. do Castelo	33
Bragança	24	Guarda	35

⁽⁵⁾ P. B. LEMA e P. M. MATHER, 1970. Citado.

⁽⁶⁾ L. M. CÂNDIDO — *Aspectos Regionais da Demografia Portuguesa*. Lisboa 1969.

O distrito interior da Guarda, as áreas da fronteira do Minho (distrito de Viana do Castelo) e as áreas montanhosas do norte da Beira (distrito de Viseu) apresentavam situações extremas de envelhecimento, enquanto Bragança e Vila Real, por terem taxas elevadas de fertilidade, não tinham atingido ainda aquela posição.

Estas considerações interessam ao problema apresentado de início — o de interpretar as características da evolução demográfica no contexto de uma diferenciação de tipo sócio-económico. Por um lado, no que diz respeito à distribuição espacial daquelas características, verifica-se que a fertilidade ou a natalidade representam um comportamento que tem pouco a ver com condicionalismos de carácter estritamente económico, enquanto outras variáveis — nomeadamente a estrutura etária — são directa e imediatamente influenciadas pelos desequilíbrios regionais «de crescimento». Mas, por outro lado, elas estão todas relacionadas num processo interactivo, de modo que as transformações têm repercussão geral, mais ou menos lenta, em todas as variáveis que no seu conjunto definem uma fase de evolução demográfica. Nesta perspectiva, facilmente se compreende que as correntes intensas de emigração, que afectaram as áreas do interior durante a última década, tivessem resultado no reforçar do contraste entre populações rurais envelhecidas e as populações mais jovens das áreas industriais. Mas, além deste padrão mais evidente, é provável que aquele factor incidisse também no comportamento da fertilidade, que tenderia a diminuir nas áreas do interior que em 1960 apresentavam valores muito próximos de concelhos dos distritos de Braga ou do Porto. Pareceu interessante desenvolver esta hipótese, utilizando toda a informação disponível.

A. A INFORMAÇÃO UTILIZADA

A análise do dinamismo demográfico em 1970 abrangeu toda a área a norte do Douro e alguns concelhos a sul deste rio. Foram excluídos o concelho do Porto e a área de migrações pendulares com a cidade — concelhos de Matosinhos, Maia, Valongo, Gondomar, Vila Nova de Gaia, Espinho — e ainda Oliveira de Azeméis, S. João da Madeira, Vila da Feira e

Ovar. O critério adoptado para a delimitação da área de influência do Porto baseou-se na análise da variação das actividades industriais e terciárias. Pareceu-nos, com base naqueles indicadores, que Estarreja estaria mais relacionada com a expansão industrial de Aveiro do que com a área metropolitana do Porto. Note-se que a delimitação adoptada nesta análise corresponde às conclusões de estudos elaborados pelo Gabinete de Planeamento da Região Norte (7).

Não puderam ser incluídas estatísticas dos movimentos da população. As relativas à emigração são deficientes; as de migração interna não existem, e, mesmo num trabalho especializado sobre esta matéria, o grau de atracção e repulsão das áreas do País teve de ser extrapolado a partir da diferença entre a população de facto e o crescimento natural (8). Houve pois que adoptar uma óptica semelhante neste trabalho: avaliar o dinamismo demográfico a partir da estrutura etária e factores naturais do crescimento da população.

Na matriz inicial incluímos a variação demográfica no decénio 1950-1960 e a estrutura etária da população em 1970. Substituímos a taxa de natalidade (número de nascimentos por 1000 habitantes) pela taxa de fertilidade (número de nascimentos por número de mulheres no grupo etário de 20 a 40 anos). A estrutura comercial não foi incluída na análise, preferindo-se uma discriminação no sector terciário das profissões liberais e empregados de escritório. Foram excluídas as variáveis que se verificou na análise anterior não terem grau elevado de correlação.

As variáveis estão indicadas na figura 5: 1 é expressa em números absolutos; 3, 4, 5, 6, 7, 8, 12, 13, 14, 15, 16 e 17 são percentagens; 2 e 11 são densidades em relação a 1000 hectares e 1000 habitantes respectivamente; 9 é o número de nascimentos por número de mulheres no grupo etário de

(7) L. V. OLIVEIRA — *Área Metropolitana do Porto — Subsídios para a sua definição*. Sec. de Planeamento Urbanístico. Centro de Estudos de Engenharia Civil. I. A. C. Fac. de Engenharia da Univ. do Porto. Maio 1973. Boletim n.º 7.

(8) A. ALARCÃO — *Mobilidade Geográfica da População de Portugal (Continente e Ilhas Adjacentes). Migrações Internas 1921-1960*. Lisboa 1969.

20 a 40 anos; 10 é a diferença entre a relação de nascimentos e a relação de mortos, ambos ponderados a 1000 habitantes.

B. ANÁLISE DE CORRELAÇÕES

Vamos analisar, em primeiro lugar, as variáveis demográficas e os indicadores sócio-económicos separadamente e, em seguida, as relações recíprocas.

Pode verificar-se na matriz (fig. 5) que a variação dos grupos etários de menos de 20 anos, de 40 a 60 anos e mais de 60 anos (5, 6, 7 e 8) está fortemente correlacionada entre si, com o crescimento natural (10) e, em menor grau, com a fertilidade (9).

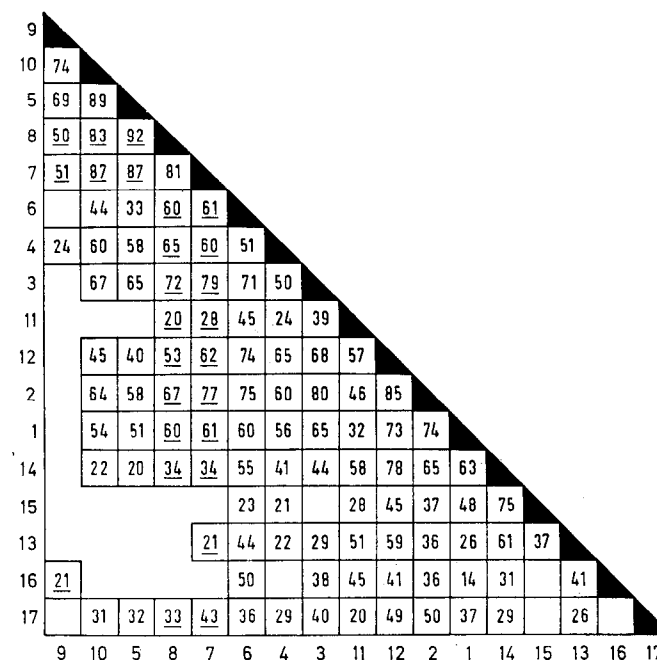
Entre os indicadores sócio-económicos são a densidade (var. 2) e a população não agrícola activa (var. 12) que apresentam maior inter-relação com as restantes; a correlação entre elas é aliás muito elevada (0,85). Seguem-se, por ordem de importância, a variação da população no último decénio e o número de efectivos (var. 3 e 1). Pelo contrário, as variáveis de transportes, profissões liberais, empregados de escritório e indústria (13, 14, 15, 16 e 17) têm correlações muito fracas. Particularmente, as variáveis 14 e 15, que representam os sectores do terciário de maior grau hierárquico, têm médias baixas e desvio padrão elevado; estão fortemente relacionadas entre si (0,75) mas correlacionam com poucas variáveis.

C. Os COMPONENTES

QUADRO III

Valor dos primeiros 4 componentes e percentagem cumulativa da variação que contém

Componente	Valor	% Cum. variação
I	8,182	48,1
II	3,103	66,4
III	1,271	73,9
IV	0,864	79,0



Variáveis:

1. População presente em 1970.
2. Densidade de população em 1970.
3. Variação da população entre 1960 e 1970.
4. Variação da população entre 1950 e 1960.
5. População com idade inferior a 20 anos.
6. População no grupo de idade de 20 a 40 anos.
7. População no grupo de idade de 40 a 60 anos.
8. População com idade superior a 60 anos.
9. Taxa de fertilidade.
10. Taxa de crescimento natural.
11. Consumo de carne por habitante.
12. População activa ocupada em actividades não agrícolas.
13. População activa ocupada nos transportes.
14. Empregados de escritório.
15. Profissões liberais.
16. População activa empregada na indústria alimentar.
17. População activa empregada na indústria têxtil.

Fig. 5 — Matriz de correlações entre variáveis. Não se consideraram os valores inferiores a 0,20.

QUADRO IV

Peso das variáveis nos dois primeiros componentes

Variáveis	1.º componente	2.º componente
1	-0,79	
2	-0,92	
3	-0,85	
4	-0,72	
5	-0,74	0,62
6	-0,77	-0,27
7	0,86	-0,35
8	0,83	-0,42
9	-0,31	0,76
10	-0,77	0,55
11	-0,50	-0,50
12	-0,87	-0,36
13	-0,43	-0,62
14	-0,67	-0,55
15	-0,40	-0,41
16	-0,37	-0,46
17	-0,51	

As variáveis com elevado grau de inter-relação com as restantes (2, 12, 7, 3, 8 e 1) têm maior peso no primeiro componente. No segundo componente a variável de maior peso é a fertilidade (9) que, como referimos, não estava relacionada com outras variáveis, além daquelas relativas aos grupos etários e ao crescimento natural (6, 7, 8 e 10). Ainda assim, a maior parte da variação destas últimas variáveis está explicada no primeiro componente.

As variáveis 13, 14, 15 e 16 têm valores baixos tanto no primeiro como no segundo componente. Elas não são totalmente explicadas por estes e aparecem ainda no terceiro componente, juntamente com a variável do grupo de idade de 20 a 40 anos, com a qual estão correlacionadas.

Observando o peso das variáveis nos componentes, podemos interpretar o seu significado. Verifica-se, assim, que o primeiro explica a tendência geral de distribuição da população: foi persistente o aumento durante as duas últimas décadas (var. 3 e 1) em áreas já densamente povoadas (var. 1 e 2) e nelas predominam os estratos mais jovens (note-se os sinais contrários das correlações das variáveis 5, 6, 7 e 8). Reflecte,

de facto, o contraste de dinamismo demográfico entre o litoral e o interior. A importância das actividades primárias é uma variável com peso elevado no componente (12 e 14), o que significa que aquele padrão demográfico se relaciona com o grau de desenvolvimento de actividades económicas. Pelo contrário, a fertilidade (var. 9) não contribui significativamente para reforçar o padrão nem sequer tem influência dominante nos contrastes de crescimento natural da população.

É no segundo componente que esta variável tem elevada correlação, a traduzir o comportamento demográfico, bem diferenciado, das áreas rurais de montanha dotadas de baixa acessibilidade e onde a fertilidade se mantém elevada.

O terceiro componente explica ainda parte da variação do grupo de idade adulta (20 a 40 anos) associada a funções urbanas (profissões liberais e empregados de escritório). Assim, enquanto o primeiro componente revela uma oposição geral entre populações jovens das áreas urbanas do litoral e populações envelhecidas das áreas rurais do interior, o terceiro componente expressa uma menor variação — a maior ocorrência de indivíduos de 20 a 40 anos nas aglomerações urbanas do litoral ou nos pequenos centros do interior. Porém, a variação daquele grupo não é totalmente explicada por aqueles componentes e aparece ainda, no quarto e no quinto, associada a outras variáveis, de modo que o significado destes componentes não coincide com o terceiro. A conclusão que se deve tirar é a de que a variabilidade do grupo de 20 a 40 anos é maior do que a dos outros grupos etários e mais difícil de interpretar. Este facto seria de esperar: a retracção do grupo de idade dos 20 a 40 anos é um facto geral da demografia do País, por ser aquele onde se fazem sentir os reflexos da emigração intensa durante os últimos decénios. A escala do País, podem estabelecer-se relações entre a intensidade do fluxo migratório e o nível sócio-económico das áreas de origem mas, numa escala mais pormenorizada, a decisão de migrar não pode ser explicada simplesmente por aquelas variáveis: por exemplo, em concelhos do distrito de Braga ocorreu uma emigração tão intensa como em alguns concelhos trasmontanos e, contudo, as condições sócio-económicas diferem entre aquelas áreas. Haveria ainda que considerar a

intensidade das migrações inter-regionais, particularmente as que têm como destino a grande área de Lisboa.

A posição dos concelhos nos 2 primeiros componentes foi reproduzida no gráfico da figura 6. Pelo que atrás foi dito, o primeiro é fácil de interpretar. Analisaremos apenas o segundo — eixo vertical — que mostra a actual diferenciação da taxa de fertilidade. O seu valor é muito baixo no sector superior do Douro — Freixo de Espada à Cinta, Moncorvo e Carrazeda de Anciães — e em outros concelhos extremos de Trás-os-Montes, que representam situações muito idênticas às da margem fronteiriça do Minho ou de áreas interiores da Guarda, onde a população já se apresentava envelhecida no decénio de 1950-1960. Verifica-se, assim, que a fertilidade não correlaciona com níveis de desenvolvimento económico mas é este último factor que desencadeia um processo com tendência para se acentuar. Nas regiões mais industrializadas mantêm-se largos estratos de população jovem, enquanto nas áreas rurais do interior, devido à intensidade de emigração, são os estratos de população idosa que predominam. Nestas últimas diminui drasticamente a taxa de fertilidade, enquanto nas áreas mais industrializadas ela estabiliza. Este padrão reflecte uma ruptura entre estruturas demográficas tradicionais por efeito do estabelecimento de uma corrente intensa e contínua de emigração. Note-se como nas áreas de montanha entre o Minho e Trás-os-Montes, onde o grau de acessibilidade e de contactos é baixo, perduram ainda taxas de fertilidade elevadas.

D. AGRUPAMENTO DE CONCELHOS

A classificação final reproduz os contrastes já enunciados (fig. 7). Destaca-se Braga como pólo industrial e centro terciário mais importante do Minho. A área de localização de indústrias abrange os concelhos vizinhos de Guimarães, Barcelos, Vila Nova de Famalicão, Vila do Conde e Santo Tirso.

A sul do Douro a indústria tem um padrão diferente. Ela localizou-se preferencialmente ao longo do eixo Lisboa-Porto. Estarreja representa uma geração recente de pólos industriais, obviamente com características muito diferentes do de Braga. Na área montanhosa contígua os reflexos daquela

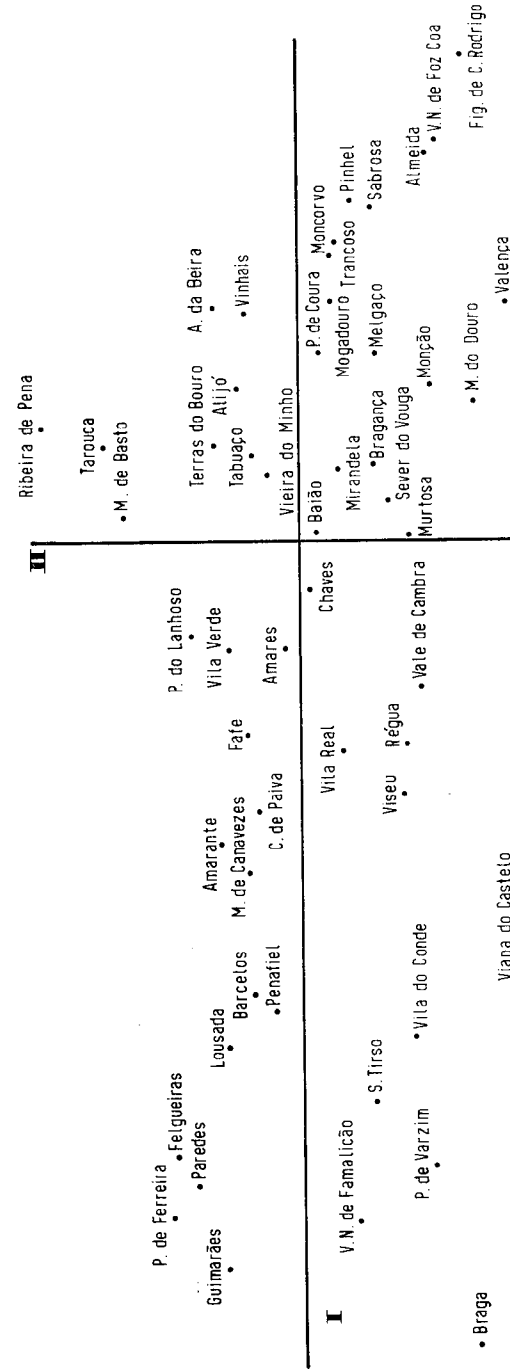


Fig. 6 — Posição dos concelhos nos dois primeiros componentes.

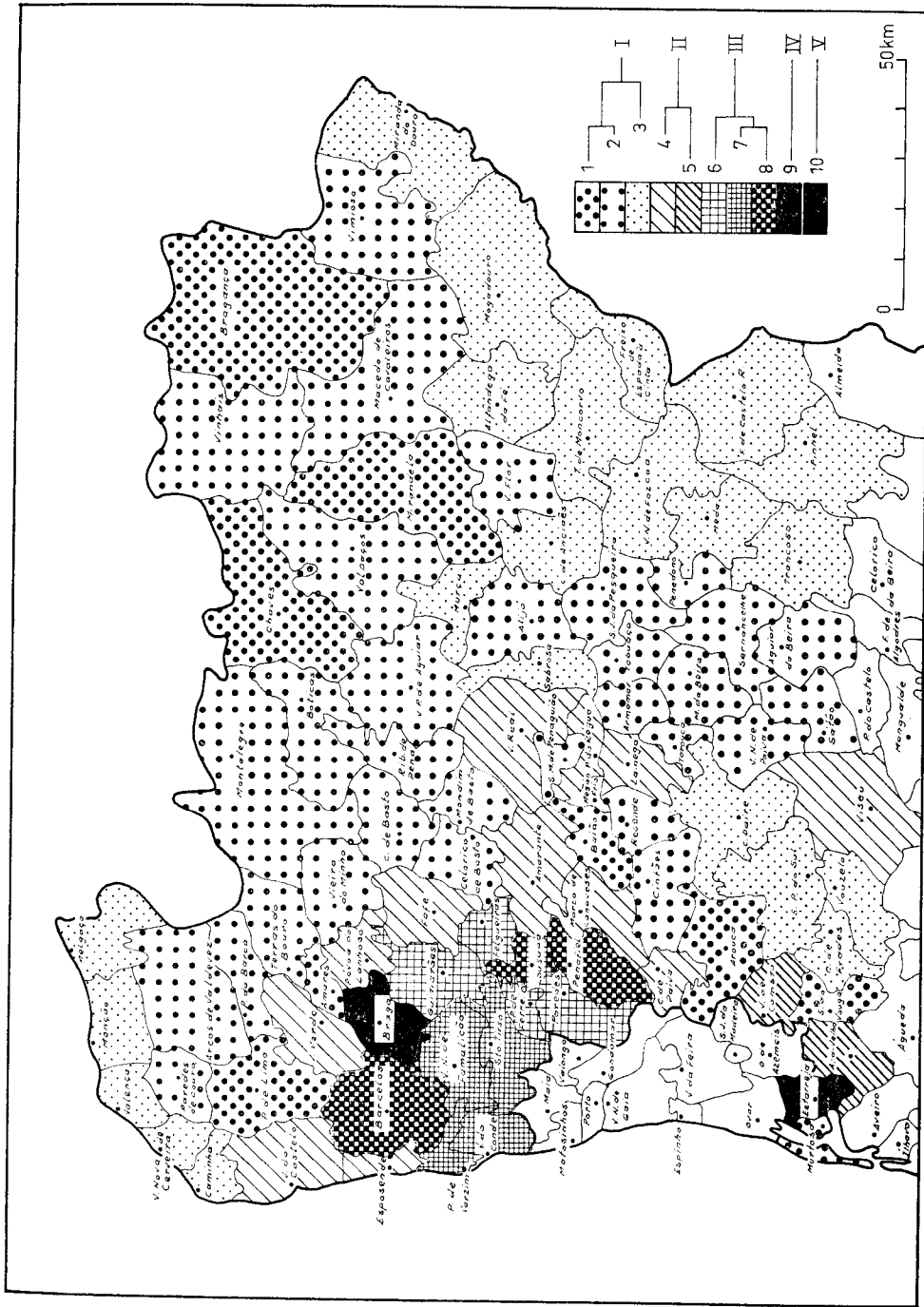


Fig. 7 — Agrupamento dos concelhos segundo o grau de semelhança.

actividade são praticamente inexistentes, exceptuando Vale de Cambra, área de indústrias agrícolas tradicionais, nomeadamente lacticínios, e onde grande parte da população é mão-de-obra operária que diariamente efectua migrações pendulares para centros industriais como S. João da Madeira e Oliveira de Azeméis. Estas condições do mercado de trabalho explicam a situação de Vale de Cambra: não apresenta semelhança com outros concelhos como Sever do Vouga e Arouca, onde o êxodo rural é mais intenso.

É interessante notar características de organização do espaço relacionadas com a implantação industrial. Ela ocorre em mancha alargada ao norte e concentrada linearmente a sul do rio Douro. Mas, ressalvando estas diferenças, é possível reconhecer como os concelhos vizinhos exteriores às áreas industriais apresentam semelhanças. Juntamente com os concelhos contíguos à área de influência dominante do Porto, constituem uma franja que é apenas interrompida pelos concelhos rurais que já apontámos — Arouca e Sever do Vouga. É de notar como Castelo de Paiva, mercê da ponte de Entre-os-Rios, que garante o acesso à margem norte do Douro, à cidade do Porto e a um centro intermediário como Penafiel, está incluída na franja. Pelo contrário, Arouca, em virtude do seu isolamento, não tem o mesmo carácter, apresentando nitidamente as características de áreas rurais de montanha, com uma economia decadente e problemas sócio-demográficos decorrentes. Já na primeira análise se tinha evidenciado uma organização espacial que concorda parcialmente com a classificação agora obtida. Por isso é preferível reservar para a parte final deste trabalho, em que se tentará relacionar os resultados das duas análises, as conclusões sobre o significado desta franja que assume características de periferia de uma área urbana e industrial.

No interior destacam-se os concelhos de Viseu, Lamego, Vila Real, Chaves, Bragança e Mirandela que correspondem à bipolarização de funções terciárias nos distritos de Viseu, Vila Real e Bragança. Dentre eles, a diferenciação de Viseu-Lamego-Régua-Vila Real justifica-se por serem importantes nós de eixos de comunicações ferroviárias e rodoviárias com funções complementares.

Contrariamente aos resultados da classificação obtida na primeira análise, as áreas do interior apresentam afinidades com as montanhas entre o Minho e Trás-os-Montes, uma vez que têm posições idênticas no componente dinamismo demográfico. A mesma razão explica a semelhança entre as áreas excêntricas e a orla fronteiriça do Minho: caracteriza-as uma evolução demográfica fortemente negativa.

IV. CONCLUSÕES

A regionalização em áreas homogéneas decorre sempre de um método de classificação que satisfaça um objectivo: simplificar as diferenças entre áreas, salientando as fundamentais. O método aqui apresentado baseia-se na análise das relações de influência recíprocas ou concorrentes entre elementos de organização espacial, e os seus resultados são eficazes para obter um agrupamento de áreas consoante o grau de semelhança nas dimensões principais. Ele representa, nesta medida, um instrumento muito importante da Geografia no capítulo da regionalização.

O método de componentes principais tem uma solução única, dependendo os resultados do tipo de informação utilizada. Ao longo deste trabalho foi insistentemente realçada a importância que assumem as características estatísticas das variáveis na aplicação do método e no produto obtido. De qualquer das análises efectuadas resultou um agrupamento de áreas homogéneas. Obtiveram-se classificações diferentes consoante as variáveis que concorrem para a discriminação entre áreas. Qualquer delas fornece um modelo espacial fácil de interpretar. Mas, além disso, a conjugação dos resultados obtidos em ambas as análises permite equacionar problemas relevantes de organização espacial. Estas conclusões resultam da potencialidade de um método matemático de solução única.

Na primeira análise, o factor mais importante de diferenciação espacial foi a localização de actividades económicas. Evidenciaram-se dois espaços — Minho e Trás-os-Montes — onde os factores de interacção são distintos, devido fundamentalmente a diferenças no desenvolvimento económico entre as duas áreas.

No Noroeste, a localização industrial condiciona a rede de funções urbanas que se organiza num contraste centro-periferia. Este efeito centro-periferia é revelado pela caracterização dos concelhos contíguos às áreas industriais de Braga e Guimarães, que apresentam maior semelhança com os centros do interior do que com aqueles aglomerados mais próximos. Na interpretação dos resultados deverá ainda ter-se em conta que as estatísticas de emprego utilizadas se referem sempre ao local de residência e não ao local de trabalho. Assim, a sua variação não expressa que as actividades se difundiram a partir de Braga para a franja contígua mas que esta corresponde a uma periferia residência de mão-de-obra operária.

Na segunda análise verifica-se que o maior dinamismo demográfico ocorre nas áreas industrializadas e diminui com a distância, delimitando claramente a área de influência dos pólos de trabalho.

Com base em elementos que aqui não foram considerados, poderá objectar-se que há indícios de um mecanismo de repercussão das estruturas urbanas e industriais nas áreas envolventes (*spread effects*): o aumento de produtividade agrícola, nomeadamente no sector pecuário, registado durante as últimas décadas no Minho, corresponde, de facto, a um alargamento de área de produção leiteira por influência dos centros de consumo ⁽⁹⁾. Mas, quanto às actividades secundárias e terciárias, não ocorreu uma dispersão; pelo contrário, houve incremento de um processo cumulativo de vantagens. Prevê-se uma maior concentração industrial em Braga e Guimarães e também a localização do terciário tenderá a concentrar-se nestas áreas, em detrimento da periferia: não apenas de funções adequadas a um pólo — como a Universidade ou o Hospital — mas também de outros serviços de nível hierárquico mas baixo.

No interior trasmontano o padrão é completamente distinto. A actividade agrícola sem condições de rentabilidade é predominante. As densidades de população são baixas e com tendência para diminuir. A rede de distribuição de bens é incipiente, localizando-se as funções de procura ocasional,

⁽⁹⁾ PAULA BORDALO LEMA — «Elementos Estatísticos. O Arrolamento de Gado de 1972». *Finisterra*, x-19, Lisboa 1975.

juntamente com as de nível hierárquico mais elevado, em pequenos centros distantes entre si. É muito recente o processo de dispersão de funções terciárias de utilização frequente, como as do ensino ⁽¹⁰⁾.

Entre o Minho e Trás-os-Montes as áreas são de montanha com uma agricultura pobre e grau de acessibilidade muito reduzido. Elas não pertencem à periferia da área industrial do Noroeste. Na realidade, a distância aos centros industriais é demasiado elevada para que estes possam constituir pólos de absorção da sua mão-de-obra, e estabeleceu-se, de preferência, uma intensa corrente de migração de grande distância para Lisboa.

A segunda análise demonstra inequivocamente as condições sócio-económicas do dinamismo demográfico. Revela a urgência de uma política adequada de aproveitamento de recursos, capaz de sustentar o êxodo da população atraída pelo litoral e o processo cumulativo que em extensas áreas do interior resultou num estado de envelhecimento extremo. As perspectivas são de agravamento, a manter-se um modelo de localização concentrada de estruturas económicas e sociais. Mas, também nas áreas que hoje ainda não revelam sintomas tão alarmantes, os problemas não podem ser subestimados. Os centros mais industrializados do Minho, se não forem objecto de correcções de estrutura, mostrarão, cada vez mais, a tendência que é já evidente: incapacidade de alargar a área de atracção. O seu dinamismo é actualmente muito inferior ao de outras áreas industriais do País, como Aveiro, Setúbal e Lisboa.

Os problemas aqui referidos, em síntese, demonstram como os métodos de análise multivariada abrem amplas perspectivas para estudos que à Geografia interessa desenvolver. A regionalização, que é o produto final do método, não assume um carácter estritamente formal e estático. Pode ser um bom ponto de partida para outras ópticas de análise — da relação entre actividades secundárias e terciárias, do poder de crescimento e difusão dos centros industriais, das forças

⁽¹⁰⁾ PAULA BORDALO LEMA — *Rede das Funções de Ensino em Trás-os-Montes. Tendências Recentes*. Lisboa. Centro de Estudos Geográficos. I. N. I. C. Estudos de Geografia Humana e Regional, n.º 4. 1976.

de interacção, das relações funcionais entre áreas. Estes campos revestem-se de grande importância, permitindo detectar os problemas estruturais e locativos da organização espacial.

É interessante confrontar os resultados das duas análises apresentadas neste trabalho com uma demarcação de áreas baseada noutro tipo de critério — o das relações funcionais.

A rede de transporte colectivo de passageiros é um dos elementos que melhor define as forças de interacção no espaço. Nesta medida, assume um significado sócio-económico relevante, o que permite utilizá-la para uma diferenciação geográfica baseada no grau de interligação ou de dependência entre áreas e centros.

Se analisarmos a evolução desta rede ao longo deste século, verifica-se que se mantém o grau de desenvolvimento diverso nos dois espaços — Minho e Trás-os-Montes — e a barreira de descontinuidade constituída pela cadeia de montanhas entreposta. Enfim, se considerarmos a organização actual da rede e a intensidade dos fluxos que articulam os espaços com os centros urbanos, chega-se a idêntica conclusão: a classificação em áreas homogéneas relaciona-se em grande parte com a demarcação de regiões funcionais. Quanto à área montanhosa, que constitui uma charneira entre o Noroeste e o Nordeste, ela teve que optar pela utilização de bens e serviços que lhe oferecem os aglomerados urbanos do litoral, uma vez que os pequenos centros do interior trasmontano têm fraca capacidade polarizadora. Assim, Mondim de Basto, no distrito de Vila Real, não tem ligações directas com este centro, o mesmo acontecendo a parte da população do concelho de Ribeira de Pena, estabelecendo-se nestes casos relações preferenciais com Guimarães. Cabeceiras de Basto está subordinada a Braga. Já Montalegre se inclui na área de influência de Chaves e a ligação a Braga tem importância secundária, apenas se justificando pela procura de bens raros ⁽¹¹⁾.

Considerando o critério de dependência funcional, as áreas de montanha que constituem a barreira entre duas regiões

⁽¹¹⁾ Estes aspectos serão desenvolvidos num trabalho sobre a organização regional de Trás-os-Montes detectada a partir dos fluxos de transporte, a publicar na série Estudos de Geografia Humana e Regional, Centro de Estudos Geográficos, I. N. I. C.

distintas estão ligadas não a Trás-os-Montes mas ao Minho, por ser esta a alternativa mais satisfatória. Porém, em relação aos centros urbanos do Noroeste, elas apresentam o grau mais baixo de acessibilidade e são, portanto, exteriores à área onde o forte grau de dependência daqueles centros se manifesta em relações intensas e frequentes. Esta conclusão já tinha sido aferida a partir dos resultados das análises multivariadas que constam deste trabalho.

PAULA BORDALO LEMA e PAUL M. MATHER

RÉSUMÉ

Le Nord de Portugal: essai d'analyse multivariée. Cet article concerne l'application de la méthode d'analyse en composantes principales, dont le but est de déterminer les facteurs fondamentaux de la variation géographique du Nord du Portugal.

Les étapes de la méthode sont d'abord indiquées. Ensuite, sont présentés les résultats des deux analyses effectuées: elles intéressent différents ensembles de variables et, dans les deux cas, la base spatiale adoptée est le canton (concelho). L'interprétation des résultats est faite dans le cadre des divisions administratives traditionnelles plus vastes des provinces: Minho au Nord-Ouest et Trás-os-Montes au Nord-Est.

Dans la première analyse la première composante, qui extrait les plus fortes corrélations parmi les variables, montre les rapports entre les tendances de la distribution de la population et le développement des activités secondaires et tertiaires. En effet, elle résume le contraste plus général entre les agglomérations industrielles du Minho et l'habitat rural de Trás-os-Montes. Indépendante de celle-ci, la deuxième composante présente une valeur plus basse; elle explique la variation de la natalité et du croît naturel (tableaux I et II).

La position relative des cantons à l'égard des composantes obtenues permet d'interpréter la différenciation en aires (figure 2). C'est la dernière étape et le but de la méthode d'analyse en composantes principales.

Si l'on considère que la proximité des points dans l'espace des composantes est la mesure de similarité entre les cantons, on peut obtenir une différenciation géographique plus précise en employant une méthode supplémentaire de classification arborescente. Le cloisonnement des cantons d'après les composantes principales aboutit à une typologie de régions homogènes (figures 3 et 4). Les cantons du Minho où se concentrent les activités industrielles et celles du secteur tertiaire supérieur sont mis en évidence. Les cantons voisins présentent plus d'affinité avec ceux de l'intérieur de Trás-os-Montes où existent les petits centres de commerce et de services. Les taux extrêmes de ruralité caractérisent les cantons de montagne entre le Minho et Trás-os-Montes.

Dans la deuxième analyse on a introduit un plus grand nombre de variables démographiques. La première composante obtenue explique la prédominance des groupes d'âges les plus jeunes dans les cantons où se localisent les activités industrielles et tertiaires (tableaux III et IV). La deuxième composante explique la variation du taux de fertilité. Cet attribut montre une tendance à la diminution dans les régions périphériques de Trás-os-Montes qui ont subi une intense émigration depuis une dizaine d'années. Les cantons montagneux du Minho où l'isolement est le plus fort maintiennent des taux élevés de fertilité, malgré le faible dynamisme qui les caractérise. Dans la figure 6 la position des cantons à l'égard des deux premières composantes est évidente.

Les résultats de la classification arborescente sont différents de ceux obtenus dans l'analyse précédente (figure 7). Le dynamisme démographique le plus fort caractérise les cantons industriels littoraux, peu étendus et situés au nord et au sud du Douro; les cantons voisins constituent une étroite frange en position intermédiaire et le vieillissement de la population est le caractère général, même dans les cantons situés à peu de distance du littoral (figure 7).

Les deux analyses permettent d'interpréter la différence d'organisation spatiale du Nord-Ouest et du Nord-Est. Le réseau de fonctions de commerce et de services du plus haut degré hiérarchique est lié à la localisation de l'industrie. Cette relation explique que les activités secondaires et tertiaires se soient concentrées en pôles comme Braga et Guimarães dans le Minho. L'aire d'influence de ceux-ci est limité aux alentours et est insignifiante dans les cantons montagneux qui constituent une barrière entre le Minho et Trás-os-Montes. Dans cette dernière région l'activité industrielle n'est pas développée et les centres de commerce et de services sont rares et dispersés.

En admettant la persistance du modèle de développement polarisé, il ne sera pas possible à la région de l'intérieur, Trás-os-Montes, d'assurer des lieux de travail suffisants pour ralentir l'exode démographique. Il ne sera pas possible non plus de disperser les services mais, bien au contraire, leur concentration dans les centres industriels du littoral ira en s'accroissant.

L'organisation spatiale du Nord de Portugal pose des problèmes très complexes. D'une part, le déséquilibre socio-économique entre le littoral et l'intérieur, au sens le plus large, c'est à dire entre le Minho et Trás-os-Montes, s'est accru, ce qui explique les forts contrastes de vieillissement de la population entre les deux régions. D'autre part, même dans le Minho, s'est aggravée l'accumulation des avantages des centres industriels au dépens de la périphérie la plus proche. Le dynamisme du Nord est déjà très faible; même dans les centres industriels il est en diminution relative par rapport à d'autres centres littoraux comme Aveiro, Lisbonne ou Setúbal. En l'absence d'un plan de réorganisation industrielle adapté à la Région Nord, la stagnation des centres industriels traditionnels va s'accroître et peu à peu, ceux-ci perdront leur fonction régionale de pôle d'attraction de main d'œuvre.

SUMMARY

Northern Portugal: an essay on a multivariable analysis. This paper aims to evaluate the basic pattern of geographic variation in Northern Portugal by means of *principal component analysis*.

In the first part the methodological context is summarized. The second and third parts present two different analysis: they differ in the variables involved and in both cases the same spatial base was used—the county. In the interpretation of results, the reference to larger administrative areas, the provinces of Minho and Trás-os-Montes, was thought more appropriate.

In the first analysis, *the first component* explains differences in population distribution related to variation of secondary and tertiary activities. In fact, it delineates the most general pattern of relationships in the data: the sharp contrast between the most urbanized areas in Minho and the rural parts of Trás-os-Montes sparsely populated. The *next greatest component*, orthogonal to the first has a smaller *eigenvalue* and summarizes mainly fertility and natural growth variation (tables I and II).

The last step of the method accounts for the individual *area scores* on the components. Thus, the final output is the significant areal differentiation according to the input, the variables initially included in the data matrix (figure 2).

A further extension is the *Cluster Analysis* (figure 3). If we assume that the distance between points in the components space is equivalent to similarity it is possible to produce a *linkage tree* such as the counties which are near are grouped in larger homogeneous areas and separated from others which are apart (heterogeneous). Figure 4 presents such a classification. The most industrialized counties in Minho have a remarkable concentration of tertiary activities of the highest level. Nevertheless, its neighbouring counties are much more similar to those outstanding in the Interior Region (Trás-os-Montes) as locations of small commerce and service centres for a rural area. The most backward counties are located in the mountains separating Minho from Trás-os-Montes.

In the second analysis a much wider range of demographic variables was used. The *first component* shows the high inter-correlation of the younger sectors of the population with industrial and tertiary activities (tables III and IV). *The second component* explains the fertility variation: this attribute is slowing down in the outmost areas due to intense migration occurred during the last decade; high rates persist in the mountain areas of Minho where accessibility is very low, yet the demographic dynamism is declining. The *area scores* on both components are represented on a graph (figure 6).

Cluster Analysis produces a classification different from that obtained in the first case: the highest demographic dynamism is confined to restrict industrial coastal areas, north and south of Douro; the bounding counties show an average situation and, as for the remaining

area, the general ageing of population is the outstanding characteristic (figure 7).

The output of both analysis contribute to evaluate the differences in spatial organization between the Northwest and the Northeast. Regarding the commerce and services, the upper levels of these functions are correlated with industrial activities and so they concentrate in poles such as Braga e Guimarães in Minho province. Its spheres of influence do not exceed the neighbouring counties and are irrelevant in the mountain areas which represent a barrier between Minho and Trás-os-Montes. In the latter region the industrial activities are almost absent, the commercial centres are few and very distant apart.

A persistent polarized developing model will maintain and enforce the migration flow from Trás-os-Montes. Furthermore, it will be unable to disperse commerce and services, on the contrary favouring its geographic concentration in the coastal industrialized centres.

The spatial organization of Northern Portugal presents acute problems: the social and economical rupture between Minho and Trás-os-Montes is aggravated which explains the strong contrasts in the demographic ageing degree between these areas; on the other hand, in Minho itself, the industrial centres show accumulated advantages at the expense of the neighbouring areas, that is to say, its nearer periphery.

Economical growth of Northern Portugal is slowing compared to other coastal areas such as Aveiro, Lisbon or Setubal. The shift encourages long distance migration towards Lisbon and other industrial areas south of Douro. This situation requires a plan aimed to the industrial reorganization of Northern Portugal; otherwise, the stagnation of its traditional centres will fasten, lessening the attractiveness to industrial and professional workers. The Region's development will be seriously handicapped and a more concentrated pattern of economic and social welfare will result.