



MAPEAMENTO DOS GASTOS EM CONSUMO DAS FAMÍLIAS DE PAÍSES DA UNIÃO EUROPEIA

(MAPPING OF HOUSEHOLD CONSUMPTION EXPENDITURE OF FAMILIES FROM EU COUNTRIES)

LUÍS N. PEREIRA

Mestre em Estatística e Gestão de Informação
Beneficiário de bolsa de investigação para doutoramento (SFRH/BD/36764/2007)
da Fundação para a Ciência e a Tecnologia
Professor Adjunto da Escola Superior de Gestão, Hotelaria e Turismo
Universidade do Algarve
lmp@ualg.pt

LARA N. FERREIRA

Mestre em Gestão e Economia da Saúde
Beneficiária de bolsa de investigação para doutoramento (SFRH/BD/25697/2005)
da Fundação para a Ciência e a Tecnologia
Professora Adjunta da Escola Superior de Gestão, Hotelaria e Turismo
Universidade do Algarve
lferrei@ualg.pt

Recebido: 06.04.2008 Aceite: 18.06.2008

RESUMO

A análise dos gastos das famílias constitui um indicador que poderá ser utilizado para estudar a performance da economia portuguesa num contexto europeu. Neste artigo apresenta-se um mapeamento da evolução dos gastos em consumo das famílias de alguns países da União Europeia, em percentagem dos gastos totais em consumo entre 1991 e 2001. Esse mapeamento é feito através de um estudo em Dupla Análise em Componentes Principais. Os resultados indicam uma evolução temporal decrescente do peso dos gastos em consumo das famílias em alimentação, bebidas, vestuário e calçado, relativamente aos gastos totais, e uma subida dos gastos em habitação, água, electricidade, gás, outros combustíveis, comunicações, entretenimento e cultura. Verificou-se a existência de uma oposição em termos de gastos em consumo entre os países do sul e os do norte da Europa. Em Portugal, observou-se uma diminuição dos gastos em consumo em alimentação, vestuário e calçado e um aumento dos gastos em entretenimento e cultura.

PALAVRAS-CHAVE

Análise em Componentes Principais; Dupla Análise em Componentes Principais; Gastos em Consumo; Mapeamento.

ABSTRACT

The analysis of the household consumption expenditure is an indicator that can be used to study the performance of the Portuguese economy in an European context. The main objective of this study is to carry out a mapping of the evolution of household consumption expenditure of some countries of the European Union, in percentage of the total consumption expenditure between 1991 and 2001. That mapping is the result of a research in Double Principal Component Analysis. The results show a decreasing tendency of the weight of household consumption expenditure in food and beverage, clothing and footwear, in proportion of total household consumption expenditure. They also show an increase of the expenditure in housing, water, electricity, gas and other fuels, communications, recreation and culture. In addition we found an opposition in terms of household consumption expenditure between the countries of the south and the ones of the north of Europe. A decrease in household consumption expenditure in food and beverage, clothing and footwear and an increase in household consumption expenditure in recreation and culture were observed in Portugal.

KEYWORDS

Principal Component Analysis; Double Principal Component Analysis; Household Consumption Expenditure; Mapping.



1. INTRODUÇÃO

A assinatura do Tratado de Roma, em 1957, ditou o início de uma nova era na Europa. De facto, a Comunidade Económica Europeia de então, constituída por seis países signatários desse Tratado, evoluiu de tal maneira que hoje em dia pouco resta daquela pequena união aduaneira. O Acto Único Europeu e, posteriormente, o tão contestado Tratado de Maastricht com os seus critérios de convergência para a criação de uma União Económica e Monetária (UEM) ditaram as bases do que hoje conhecemos como União Europeia (UE) (Silva, 1997). De meia dúzia de países para os actuais 27 Estados-membro, esta Europa renovada traduz-se numa nova realidade, a união monetária europeia, sendo esta uma provável rampa de lançamento para uma futura Europa política.

O caminho percorrido por Portugal desde a sua entrada para a UE foi particularmente impressionante, uma vez que quando a ela aderiu, a sua economia estava marcada por importantes desequilíbrios estruturais, e, embora a sua situação actual careça ainda de correcções inevitáveis, a situação sofreu mudanças incontestáveis a fim de tornar realidade uma verdadeira convergência europeia. Ainda que estas mudanças se devam em grande parte às políticas seguidas pelas autoridades portuguesas desde 1985, importa sublinhar o impacto do processo de convergência real iniciado a nível europeu desde a assinatura do Tratado de Maastricht e o papel favorável e estabilizador desempenhado pelo Sistema Monetário Europeu (Silva, 1997). É neste contexto que adquire relevância a análise de alguns indicadores relativos à performance da economia portuguesa num contexto europeu. Mais especificamente, os gastos das famílias constituem um indicador que dispensa apresentações acerca da sua importância como revelador da situação económica dos países.

Neste trabalho pretende-se realizar um mapeamento da evolução dos gastos em consumo das famílias de alguns países da UE, em percentagem dos gastos totais em consumo, no período imediatamente anterior à entrada em vigor do Tratado de Maastricht e no pós-Tratado de Maastricht, de 1991 a 2001. O mapeamento da evolução dos gastos em consumo das famílias da Dinamarca, Alemanha, França, Irlanda, Itália, Holanda, Áustria, Finlândia, Reino Unido e Portugal é feito através de um estudo em Dupla Análise em Componentes Principais. Em primeiro lugar faz-se uma análise da evolução global ao longo do tempo, em seguida efectua-se uma

análise das onze nuvens de países, e por último, analisa-se a trajectória de cada um dos países ao longo do tempo.

2. METODOLOGIA

Foram utilizados dados secundários disponíveis no sítio da Internet do EUROSTAT (*Statistical Office of the European Communities*) (EUROSTAT, s.d.) para um conjunto de rubricas referentes ao período entre 1991 e 2001. As rubricas consideradas na estrutura dos gastos em consumo das famílias da UE são as seguintes: X1-Alimentação e bebidas não alcoólicas, X2-Bebidas alcoólicas, tabaco e narcóticos, X3-Educação, X4-Vestuário e calçado, X5-Habituação, água, electricidade, gás e outros combustíveis, X6-Comunicações, X7-Entretenimento e cultura, X8-Saúde. A análise de dados foi efectuada no programa estatístico Statistical Analysis System (SAS) versão 9.13 e no Microsoft Excel 2003.

2.1. Dupla Análise em Componentes Principais

A Dupla Análise em Componentes Principais (DACP) é uma técnica de análise de dados desenvolvida por Bouroche (1975) adequada para analisar dados cúbicos. A DACP é utilizada para situações em que um mesmo conjunto de p variáveis é avaliado para um mesmo conjunto de n indivíduos em T momentos do tempo, como por exemplo séries temporais. O objectivo global da DACP consiste em comparar globalmente a evolução da associação entre as variáveis do estudo e a evolução dos indivíduos ao longo dos instantes $1, 2, \dots, T$. Pelo facto da bibliografia sobre esta metodologia ser parca, são em seguida resumidos os seus princípios básicos, fundamentais para a dedução das ajudas clássicas à interpretação, remetendo-se o leitor para as obras Bouroche (1975), Bouroche e Dussaix (1975), Kroonenberg (1983), Lavit (1988) e Ambapour (2001). É de referir que este artigo é um dos primeiros da literatura no qual são apresentadas as ajudas clássicas à interpretação no contexto da DACP. Algumas referências detalhadas sobre Análise em Componentes Principais (ACP), fundamentais para a DACP, podem ser encontradas, por exemplo, em Bouroche e Saporta (2002), Lebart, Morineau e Piron (2000), Johnson e Wichern (2002) e Saporta (1990).

Designa-se o número de indivíduos em estudo por n , o número de variáveis por p (que se supõem centradas e reduzidas) e o número de períodos de tempo por T . O quadro completo de dados, X , de



dimensão $(nT \times p)$ é dado por: $X = col_{1 \leq t \leq T}(X_t)$ onde $X_t = col_{1 \leq i \leq n} \begin{pmatrix} x_i^{(t)} \\ \vdots \\ x_i^{(t)} \end{pmatrix}$ é uma matriz de dimensão $(n \times p)$ com indivíduos $x_i^{(t)} = col_{1 \leq j \leq p} [x_i^{(t)}]$. No momento t , a variável j representa-se por $(x_i^{(t)})^j = col_{1 \leq i \leq n} [(x_i^{(t)})^j]$, onde $(x_i^{(t)})^j$ representa o valor do indivíduo i , na variável j no momento t . No plano teórico, conclui-se então que DACP tem como objectivo a comparação de T quadros de dados que podem ser representados por um tríplice (X_t, Q, D) , onde Q é a métrica no espaço dos indivíduos e D é a métrica no espaço das variáveis. Acresce ainda que a DACP permite calcular medidas globais de relação entre os quadros de dados (estudos), baseadas no coeficiente RV de Escoufier (1973). A relação entre os trílices (X_k, Q, D) e (X_j, Q, D) é dado pelo coeficiente de correlação vectorial:

$$RV(X_k, X_j) = Cov(X_k, X_j) [Var(X_k) Var(X_j)]^{-1/2}, \quad (1)$$

onde $Cov(X_k, X_j) = tr(X_k Q X_k' D X_j Q X_j' D)$

e $Var(X_k) = tr(X_k Q X_k' D)^2$. O coeficiente RV pode ser interpretado como a concordância da estrutura de co-inércia de dois quadros de dados produzida pela sua estrutura de inércia. O seu valor está compreendido entre zero (não existe relação entre as variáveis dos grupos considerados) e um (as nuvens que representam os grupos são homotéticas).

A DACP é efectuada em três etapas. Na primeira etapa é efectuada uma análise da evolução global ao longo do tempo, na segunda etapa é efectuada uma análise das T nuvens de indivíduos e na terceira etapa é efectuado um estudo da intraestrutura. A “análise da evolução global ao longo do tempo” consiste em realizar uma ACP dos centros de gravidade das nuvens (estudo da interestrutura). Esta análise conduz a uma imagem euclidiana dos quadros num espaço de dimensão conveniente. De um modo geral, o primeiro eixo principal de inércia desta imagem explicita a evolução dos centros de gravidade ao longo do tempo. A “análise das T nuvens de indivíduos” consiste em efectuar uma ACP a cada um dos T quadros de dados, X_t , o que permite interpretar os resultados de cada ACP ao nível dos indivíduos e ao nível das variáveis. O “estudo da intraestrutura” consiste na determinação de um subespaço comum de representação das T nuvens de indivíduos. Neste estudo pretende-se caracterizar o subespaço de projecção, gerado pela base de vectores $\begin{pmatrix} \varphi \\ \vdots \\ \varphi \end{pmatrix}_{K=1,2,\dots,q}$ no qual será possível representar as trajectórias dos indivíduos ao longo do tempo, com base num critério de optimização a definir. Uma vez que as duas primeiras etapas da

DACP consistem na realização de simples ACP, neste artigo são apenas apresentados resultados detalhados respeitantes ao “estudo da intraestrutura”, bem como desenvolvimentos ao nível das ajudas à interpretação clássicas de uma análise factorial para esse estudo.

A selecção do “melhor” sistema de eixos $\begin{pmatrix} \varphi \\ \vdots \\ \varphi \end{pmatrix}_{K=1,2,\dots,q}$,

no âmbito do “estudo da intraestrutura”, pode ser determinada segundo três critérios: (i) minimização da soma das perdas de inércia, (ii) procura sequencial do novo sistema de eixos e (iii) maximização da inércia explicada. A comparação da performance de soluções, produzidas por diferentes critérios, pode ser efectuada através de um índice. Para um determinado sistema de eixos,

$$\begin{pmatrix} \varphi \\ \vdots \\ \varphi \end{pmatrix}_{K=1,2,\dots,q}, \text{ o índice} \quad (2)$$

$$\Phi(t, \varphi) = \left[\sum_{k=1}^q \lambda_k^{(t)} - \sum_{k=1}^q \varphi_k^T V_t \varphi_k \right] \times \left[\sum_{k=1}^q \lambda_k^{(t)} \right]^{-1}$$

mede a perda relativa de inércia da nuvem $N(t)$ dos indivíduos associados ao quadro t , quando os mesmos são projectados num subespaço gerado pela

base de vectores $\begin{pmatrix} \varphi \\ \vdots \\ \varphi \end{pmatrix}_{K=1,2,\dots,q}$, em vez de serem projectados no subespaço associado aos q primeiros eixos. Note-se que a inércia da nuvem $N(t)$ explicada pelo sistema $\begin{pmatrix} \varphi \\ \vdots \\ \varphi \end{pmatrix}_{K=1,2,\dots,q}$ é dada por $\sum_{k=1}^q \varphi_k^T V_t \varphi_k$.

Segundo o critério da minimização da soma das perdas de inércia, a soma das perdas de inércia quando se projectam os indivíduos de cada uma das nuvens $N(t)$ ($t=1, 2, \dots, T$), no subespaço gerado por esse sistema de eixos, deve ser mínima. Desta forma, a solução óptima $\begin{pmatrix} \varphi \\ \vdots \\ \varphi \end{pmatrix}_{K=1,2,\dots,q}$ é tal que $\Phi(\cdot, \varphi) = \frac{1}{T} \sum_{t=1}^T \Phi(t, \varphi)$

seja mínimo. Alternativamente pode recorrer-se à maximização da quantidade $\sum_{t=1}^T \sum_{k=1}^q \cos^2 \left(\varphi, z_{-k}^{(t)} \right)$.

Estes problemas de optimização podem ser facilmente resolvidos com o apoio informático.

A procura sequencial do novo sistema de eixos é efectuada iterativamente. Na primeira iteração determina-se o vector φ_{-1} tal que $\sum_{t=1}^T \cos^2 \left(\varphi_{-1}, z_{-1}^{(t)} \right)$ seja máximo; na segunda iteração determina-se o vector φ_{-2} , ortogonal a φ_{-1} , tal que $\sum_{t=1}^T \cos^2 \left(\varphi_{-2}, z_{-2}^{(t)} \right)$ seja máximo; na e -ésima iteração determina-se o vector φ_{-e} , ortogonal ao subespaço gerado por $\varphi_{-1}, \dots, \varphi_{-e-1}$, tal que $\sum_{t=1}^T \cos^2 \left(\varphi_{-e}, z_{-e}^{(t)} \right)$ seja máximo; e assim sucessivamente



até à construção do vector φ_{-q} .

Por último, segundo o critério da maximização da inércia explicada, pretende-se maximizar a inércia do conjunto das T nuvens projectadas:

$$\text{Max}_{\varphi_{-1}, \dots, \varphi_{-q}} \sum_{t=1}^T \sum_{k=1}^q \varphi_{-k}^T V_t \varphi_{-k} \quad (3)$$

Admitindo-se que a influência das diferentes nuvens $N(t)$ dos indivíduos na nova imagem euclidiana é equilibrada, pode-se considerar que $V = \sum_{t=1}^T V_t$. Nesta

situação, o problema de maximização (3) é equivalente a $\text{Max}_{\varphi_{-1}, \dots, \varphi_{-q}} \sum_{k=1}^q \varphi_{-k}^T V \varphi_{-k}$. A solução deste problema é uma conhecida solução da ACP: os vectores $\begin{pmatrix} \varphi_{-k} \end{pmatrix}_{k=1,2,\dots,q}$ são os q vectores próprios da matriz $QV = \sum_{t=1}^T QV_t$.

Este critério consiste em efectuar uma ACP da nuvem dos “ nT indivíduos” centrados em relação ao seu centro de gravidade (para cada um dos instantes t) e definidos pelas p variáveis do estudo.

No espaço das variáveis, admite-se que todos os indivíduos têm o mesmo peso definindo-se $D = \frac{1}{n} I_n$.

A métrica utilizada no espaço dos indivíduos é $Q = I_p$ porque se admite que os quadros X_t , estão centrados e reduzidos. Neste caso particular, os vectores

$\begin{pmatrix} \varphi_{-k} \end{pmatrix}_{k=1,2,\dots,q}$ são os q vectores próprios da matriz $V = \sum_{t=1}^T V_t$, onde $V_t = X_t^T D X_t$ é a matriz de variância-covariâncias associada à nuvem $N(t)$.

As coordenadas dos pontos das T nuvens sobre o k -ésimo eixo principal são dadas por $\underline{Y}^k = X \varphi_{-k}$, sendo as coordenadas dos pontos da nuvem $N(t)$

dadas por $(\underline{Y}^k)^{(t)} = X_t \varphi_{-k}$. A k -ésima componente principal associada ao período t apresenta média nula e variância igual a $\varphi_{-k}^T V_t \varphi_{-k}$. A partir dos resultados anteriores, e considerando os resultados obtidos na ACP, obtém-se facilmente as ajudas à interpretação clássicas, ao nível dos indivíduos e ao nível das variáveis. Em primeiro lugar, ao nível dos indivíduos: a contribuição do i -ésimo indivíduo para a formação do eixo k no momento t é dada por

$(CTA_i^k)^{(t)} = p_i \left[(y_i^k)^{(t)} \right]^2 \times \left[\varphi_{-k}^T V_t \varphi_{-k} \right]^{-1}$, a contribuição global do i -ésimo indivíduo para a formação do eixo k

é dada por $CTA_i^k = \sum_{t=1}^T p_i \left[(y_i^k)^{(t)} \right]^2 \times \left[\sum_{t=1}^T \varphi_{-k}^T V_t \varphi_{-k} \right]^{-1}$, e a qualidade de representação do i -ésimo indivíduo

sobre o eixo k no momento t é dada por

$(CTR_i^k)^{(t)} = \left[(y_i^k)^{(t)} \right]^2 \times \left[\varphi_{-k}^T X_t^{(i)} X_t^{(i)} \varphi_{-k} \right]^{-1}$. Ao nível das variáveis, tem-se que a contribuição da j -ésima variável para a formação do eixo k é dada por $CTA_j^k = (\varphi_k^j)^2$ e a sua qualidade de representação sobre o referido eixo é dada por $CTR_j^k = \left[\lambda_k (\varphi_k^j)^2 \right] \times \left[Ts^2 (x^j) \right]^{-1}$.

3. RESULTADOS

3.1. Análise da evolução global ao longo do tempo

Em primeiro lugar foi feita uma análise da evolução global ao longo do tempo, ou seja, uma ACP dos centros de gravidade associados a cada nuvem de indivíduos $N(t)$, $t = 1991, \dots, 2001$. A análise da tabela de dados dos centros de gravidade (médias) das 11 nuvens de países começou por ser efectuada com o cálculo das médias e dos desvios-padrão das variáveis em estudo. A observação destes indicadores permitiu verificar que os dados não são homogéneos, tendo sido necessário aplicar uma ACP normada para que todas as variáveis tivessem a mesma importância. No espaço das variáveis, admitiu-se que todos os anos têm o mesmo peso.

O cálculo dos coeficientes RV (Tabela 1, página seguinte) permite observar que existe uma forte correlação entre os vários pares de quadros de dados referentes aos períodos 1991-1994 e 1997-2001, o que significa que nesses períodos a estrutura dos diferentes quadros de dados não evoluiu de forma significativa. Os anos de 1995 e 1996 apresentam uma estrutura diferente dos restantes, o que poderá indicar que esses são pontos de viragem.

Os dois primeiros eixos principais obtidos no estudo da interestrutura explicam 91,7% da inércia total. Uma vez que esses dois eixos explicam mais de 80% da inércia total, e porque os valores próprios associados são superiores a uma unidade, toda a análise da evolução global ao longo do tempo foi baseada apenas nos dois primeiros eixos.

A análise da tabela 2 permite verificar que a formação do primeiro eixo é devida, sobretudo, à contribuição dos anos 1991, 1992, 1993, 2000 e 2001, e de todas as variáveis com excepção das variáveis Educação (X_3) e Saúde (X_8), pelo facto de apresentarem correlações fortes, tal como se pode observar pelo gráfico 1. No primeiro eixo, que pode ser interpretado como um “factor tempo”, verifica-se



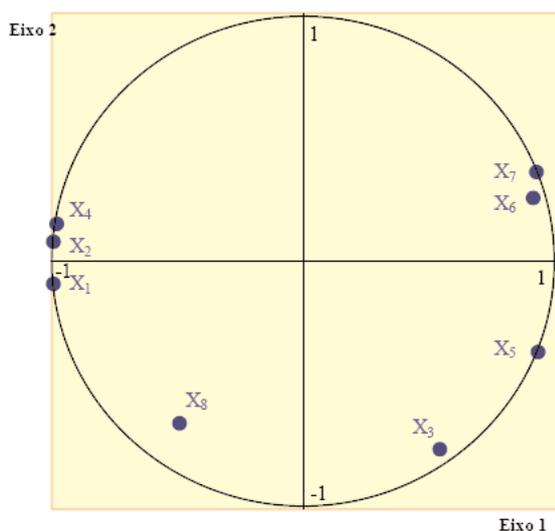
Tabela 1 – Coeficientes RV

Anos	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
1991	1,00	0,99	0,98	0,97	0,93	0,90	0,91	0,90	0,88	0,89	0,89
1992	0,99	1,00	0,99	0,98	0,93	0,89	0,90	0,89	0,87	0,87	0,87
1993	0,98	0,99	1,00	0,99	0,95	0,91	0,90	0,88	0,86	0,87	0,87
1994	0,97	0,98	0,99	1,00	0,96	0,91	0,90	0,89	0,87	0,87	0,87
1995	0,93	0,93	0,95	0,96	1,00	0,96	0,94	0,93	0,90	0,91	0,90
1996	0,90	0,89	0,91	0,91	0,96	1,00	0,98	0,96	0,93	0,94	0,93
1997	0,91	0,90	0,90	0,90	0,94	0,98	1,00	0,99	0,98	0,98	0,97
1998	0,90	0,89	0,88	0,89	0,93	0,96	0,99	1,00	0,99	0,99	0,98
1999	0,88	0,87	0,86	0,87	0,90	0,93	0,98	0,99	1,00	0,99	0,98
2000	0,89	0,87	0,87	0,87	0,91	0,94	0,98	0,99	0,99	1,00	0,99
2001	0,89	0,87	0,87	0,87	0,90	0,93	0,97	0,98	0,98	0,99	1,00

Tabela 2 – Componentes principais normadas ao valor próprio e ajudas à interpretação na ACP dos centros de gravidade

Anos	Eixo 1			Eixo 2			
	Y ¹	CTA	CTR	Y ²	CTA	CTR	Qualidade
1991	-3,97	24,04	71,47	2,45	40,62	27,74	99,21
1992	-3,28	16,36	97,21	0,46	1,40	1,92	99,12
1993	-2,38	8,63	71,58	-1,28	10,89	20,74	92,32
1994	-2,12	6,82	54,85	-1,50	14,86	27,45	82,30
1995	-0,19	0,05	1,39	-0,76	3,83	22,86	24,25
1996	0,70	0,74	15,26	-1,33	11,72	55,17	70,43
1997	1,08	1,78	57,56	-0,71	3,34	24,86	82,43
1998	1,65	4,16	81,97	0,50	1,65	7,46	89,42
1999	2,32	8,19	84,73	0,87	4,98	11,84	96,56
2000	2,92	12,96	86,08	0,95	6,00	9,16	95,24
2001	3,27	16,27	93,01	0,33	0,70	0,93	93,94

Gráfico 1 – Círculo de correlações na ACP dos centros de gravidade



- X₁-Alimentação e bebidas não alcoólicas
- X₂-Bebidas alcoólicas, tabaco e narcóticos
- X₃-Educação
- X₄-Vestuário e calçado
- X₅-Habitação, água, electricidade, gás e outros combustíveis
- X₆-Comunicações
- X₇-Entretenimento e cultura
- X₈-Saúde

uma oposição entre os anos mais longínquos e os anos mais recentes. Através de uma análise conjunta dos resultados expostos na tabela 2 e no gráfico 1, observa-se, ao longo do período 1991-2001, uma evolução temporal decrescente *quasi*-linear dos gastos em consumo das famílias em Alimentação e bebidas não alcoólicas (X₁), Bebidas alcoólicas, tabaco

e narcóticos (X₂) e Vestuário e calçado (X₄), em percentagem dos gastos totais em consumo. Por outro lado, durante esse mesmo período observa-se uma evolução temporal crescente dos gastos em consumo das famílias em Habitação, água, electricidade, gás e outros combustíveis (X₅), Comunicações (X₆) e Entretenimento e cultura (X₇).



No que se refere ao segundo eixo principal de inércia, os anos 1991, 1993, 1994 e 1996 (Tabela 2), e as variáveis Educação (X_3) e Saúde (X_8) (Gráfico 1), têm uma acção determinante para a sua formação. Neste eixo, que pode ser interpretado como “gastos das famílias em educação e em saúde”, verifica-se uma oposição entre o ano 1991 e os anos 1993, 1994 e 1996, tendo os gastos em consumo das famílias, em percentagem dos gastos totais em consumo, em Educação (X_3) e em Saúde (X_8) apresentado valores máximos (globais e locais, respectivamente) no ano de 1996.

A representação da interestrutura evidencia “clusters de anos” que são semelhantes. Assim, conjugando a análise dos gráficos 1 e 2 observa-se o seguinte:

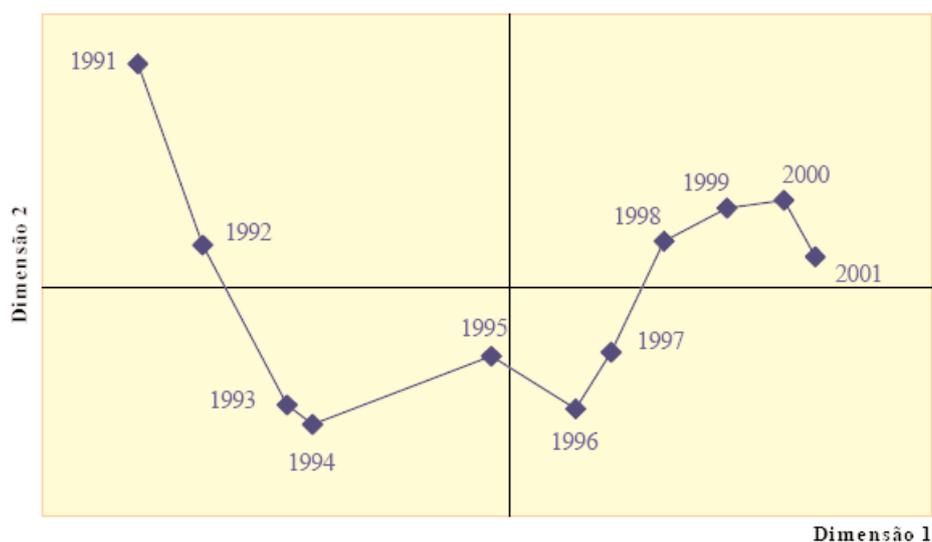
- Os primeiros quatro anos (1991-1994) são caracterizados por evoluções lineares fortes nos gastos em consumo das famílias da UE em algumas rubricas. Neste período, os gastos das famílias em consumo de Educação (X_3), Habitação, água, electricidade, gás e outros combustíveis (X_5) e em Saúde (X_8) são crescentes, por oposição aos seus gastos em Alimentação e bebidas não alcoólicas (X_1), Bebidas alcoólicas, tabaco e

narcóticos (X_2) e Vestuário e calçado (X_4), que são decrescentes;

- O período 1995-1998 é caracterizado sobretudo pela evolução das variáveis mais correlacionadas com a segunda componente principal. Desta forma, no início deste período observa-se uma estagnação e um decréscimo dos gastos em consumo das famílias em Educação (X_3) e em Saúde (X_8), seguindo-se um crescimento nos gastos em consumo nestas rubricas que atinge o valor máximo em 1996. Nos últimos dois anos deste período, o consumo das famílias da UE em percentagem dos gastos totais em consumo, em Educação (X_3) e em Saúde (X_8) volta a decrescer;

- Os últimos três anos (1999-2001) são caracterizados, inicialmente por uma estagnação, seguida de uma inversão da evolução nos gastos em consumo das famílias em algumas rubricas. A partir do ano 2000, por um lado começa a observar-se um crescimento nos gastos em consumo, em percentagem dos gastos totais em consumo, em Alimentação e bebidas não alcoólicas (X_1), Educação (X_3) e Habitação, água, electricidade, gás e outros combustíveis (X_5); enquanto por outro lado, observa-se um decréscimo nos gastos em Entretenimento e cultura (X_7).

Gráfico 2 – Representação da evolução dos centros de gravidade ao longo do tempo



3.2. Análise das T nuvens de países

Nesta etapa da DACP, é feita uma ACP normada de cada uma das 11 nuvens de países. Os resultados de cada uma destas análises conduziram à retenção, em cada análise, dos três primeiros eixos principais, que explicam cerca de 80% da inércia total. Em seguida faz-se uma breve interpretação dos resultados obtidos em cada uma dessas análises. Para simplificar a aná-

lise, decidiu-se interpretar a projecção dos dez países em cada um dos primeiros planos principais.

No período 1991-1995, os países que mais contribuem para a formação do primeiro eixo de cada uma das análises são Portugal, Finlândia, Itália e Dinamarca, observando-se uma oposição entre os países do sul e



os do norte da Europa. Nesse mesmo período, as variáveis Educação (X_3), Vestuário e calçado (X_4), Habitação, água, electricidade, gás e outros combustíveis (X_5) e Entretenimento e cultura (X_7) são aquelas que têm um maior contributo para a formação do primeiro eixo. A partir de 1996, no primeiro eixo de cada uma das análises passa a observar-se uma oposição apenas entre a Dinamarca e os dois países do sul da Europa (Portugal e Itália).

Ao longo de todo o período 1991-2001, Portugal e Itália são os países onde os gastos em consumo das famílias em Vestuário e calçado (X_4) são os mais elevados, e em Habitação, água, electricidade, gás e outros combustíveis (X_5) e Entretenimento e cultura (X_7) são os mais baixos. As famílias destes países são também caracterizadas por apresentarem os maiores gastos em consumo em Educação (X_3) no período 1991-1995, e os maiores gastos em consumo em Alimentação e bebidas não alcoólicas (X_1) a partir de 1996. Os primeiros eixos dos anos 1991-1995 exprimem ainda que os gastos em consumo das famílias finlandesas em Habitação, água, electricidade, gás e outros combustíveis (X_5) e Entretenimento e cultura (X_7) são os mais elevados, enquanto os gastos em consumo em Alimentação e bebidas não alcoólicas (X_1) e Vestuário e calçado (X_4) são os mais baixos. Por último, os gastos em consumo das famílias dinamarquesas são caracterizadas por serem os maiores nas rubricas Habitação, água, electricidade, gás e outros combustíveis (X_5) e Entretenimento e cultura (X_7) e os menores na rubrica Vestuário e calçado (X_4), ao longo dos anos 1991-2001.

No que se refere ao segundo eixo principal de inércia, os países que mais contribuem para a sua formação são a Finlândia, a Irlanda e o Reino Unido até ao ano de 1995; e a Holanda, a Irlanda e o Reino Unido nos anos seguintes. No período 1991-1995 existe uma oposição entre a Irlanda e o Reino Unido, que deixa de existir no período seguinte.

As mensagens transmitidas pelos segundos eixos principais de inércia de cada uma das análises apresentam também um ponto de viragem no ano de 1996. Desta forma:

- No período 1991-1995, em oposição aos gastos em consumo das famílias britânicas, os gastos das famílias irlandesas e finlandesas em Alimentação e bebidas não alcoólicas (X_1) e em Bebidas alcoólicas, tabaco e narcóticos (X_2) são os mais elevados, enquanto os gastos em consumo em Comunicações (X_6) são os mais baixos;

- No período 1996-2001, os gastos em consumo das famílias irlandesas e britânicas em Educação (X_3) são os mais elevados, enquanto os gastos em consumo em Comunicações (X_6) e Saúde (X_8) são os mais baixos, ao contrário do que se verifica com os gastos das famílias holandesas.

3.3. Representação das trajectórias dos países ao longo do tempo

Nesta última etapa da DACP, procurou-se um espaço comum de representação das 11 nuvens de países, no qual é possível representar as trajectórias dos 10 países ao longo do período 1991-2001. Com base no critério da maximização da inércia explicada, foi seleccionado um sistema de três eixos, que explicam cerca de 81% da inércia total.

A perda relativa de inércia de cada uma das nuvens de países $N(t)$, $t = 1991, \dots, 2001$, quando estes são projectados num subespaço tridimensional, em vez de no subespaço associado aos três primeiros eixos, é sempre inferior a 6,3% (Tabela 3). Na realidade, com o sistema de eixos seleccionado segundo o critério da maximização da inércia explicada, a perda média de inércia é menor quando se faz a projecção dos países de cada uma das nuvens no subespaço tridimensional (perda média de inércia igual a 3,1%), do que quando se faz essa projecção num subespaço bidimensional (perda média de inércia igual a 7,6%).

Tabela 3 - Perda relativa de inércia de cada nuvem de países

Nuvem	Subespaço bidimensional	Subespaço tridimensional
1991	3,5%	6,3%
1992	3,4%	5,9%
1993	2,1%	3,3%
1994	2,5%	4,9%
1995	0,8%	1,9%
1996	6,6%	0,6%
1997	9,6%	1,0%
1998	9,9%	2,0%
1999	13,1%	2,8%
2000	14,2%	2,0%
2001	17,9%	3,4%

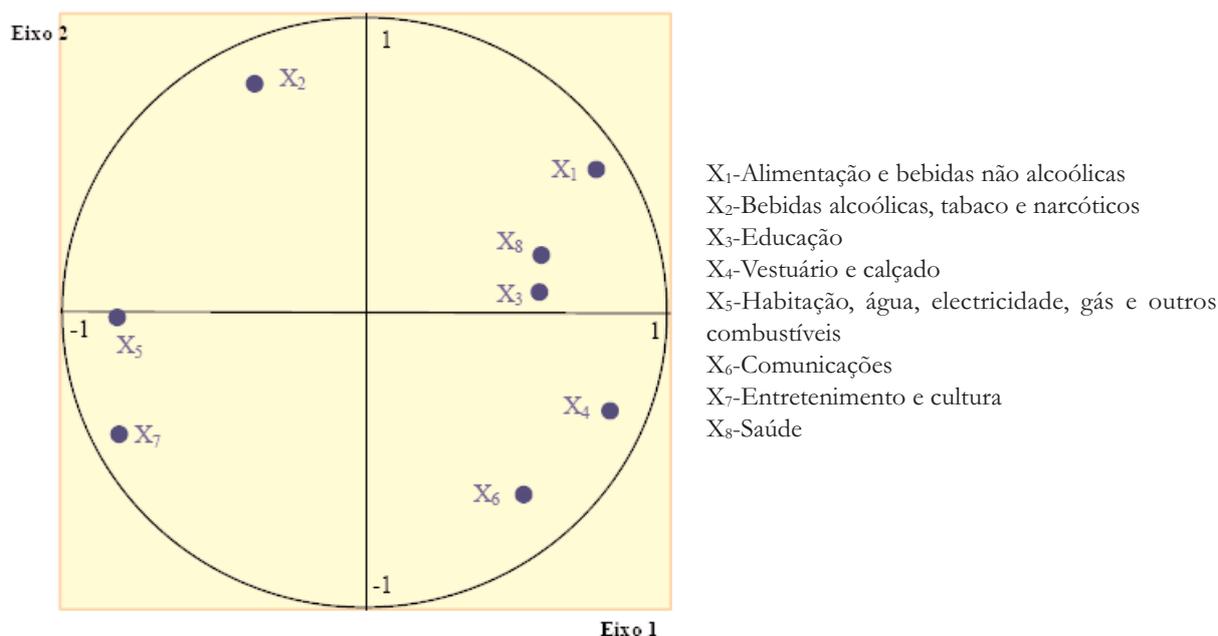
A interpretação dos eixos é possível através da análise das correlações entre as componentes principais e as “variáveis-compromisso”. As variáveis que têm uma contribuição determinante para a formação do primeiro eixo são: Alimentação e bebidas não alcoólicas (X_1), Vestuário e calçado (X_4), Habitação, água, electricidade, gás e outros combustíveis (X_5) e Entretenimento e cultura (X_7). Todas estas variáveis estão fortemente correlacionadas com a primeira



componente principal, verificando-se, a partir do gráfico 3, uma oposição entre as variáveis Alimentação e bebidas não alcoólicas (X_1) e Vestuário e calçado (X_4), e as variáveis Habitação, água, electricidade, gás e outros combustíveis (X_5) e Entretenimento e cultura (X_7). O segundo eixo opõe essencialmente a variável Bebidas alcoólicas, tabaco e narcóticos (X_2) à variável

Comunicações (X_6) (Gráfico 3). O terceiro eixo reflecte uma oposição entre as variáveis relacionadas com os gastos em consumo das famílias em educação e em saúde. A Educação (X_3) está correlacionada negativamente com a terceira componente principal, enquanto a Saúde (X_8) está correlacionada positivamente com esta componente principal.

Gráfico 3 – Círculo de correlações do estudo da intraestrutura

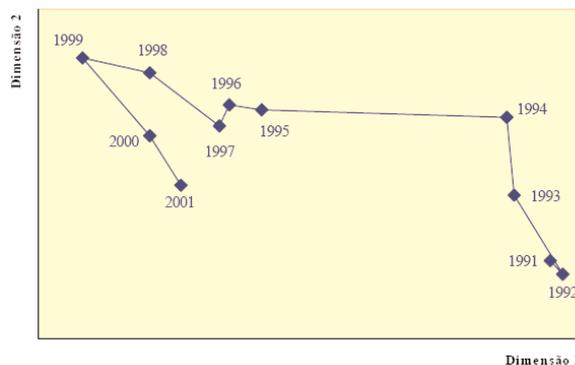


Em seguida apresentam-se as representações das trajectórias dos dez países da UE ao longo do tempo num subespaço comum bidimensional. Apesar de ter sido seleccionado um sistema de três eixos, decidiu-se apresentar neste texto estas representações de trajectórias num subespaço bidimensional para facilitar a sua interpretação. Desta forma, nos gráficos seguintes apresentam-se as trajectórias dos sete países da UE que têm uma contribuição determinante para a formação dos dois primeiros eixos: Dinamarca, Irlanda, Itália, Holanda, Áustria, Portugal e Finlândia. Apenas as trajectórias da Dinamarca e de Portugal no período 1991-2001 estão bem representadas no subespaço comum bidimensional. Nesse mesmo subespaço, estão mal representadas as trajectórias da Finlândia no período 1997-2001, da Irlanda no período 1996-2001 e da Holanda nos anos 1991, 1992, 1993, 1996 e 1997.

Em Portugal, o período 1992-1994 caracteriza-se por um crescimento dos gastos em consumo das famílias em Bebidas alcoólicas, tabaco e narcóticos (X_2). No período 1994-1999 observa-se uma evolução decrescente *quasi*-linear dos gastos em consumo em Alimentação e bebidas não alcoólicas (X_1) e Vestuário e

calçado (X_4), e uma evolução crescente, também *quasi*-linear, nos gastos em Entretenimento e cultura (X_7) (Gráfico 4). Pelo contrário, a partir do ano de 1999 observa-se um crescimento acentuado dos gastos em consumo das famílias portuguesas em Alimentação e bebidas não alcoólicas (X_1) e em Comunicações (X_6), e um decréscimo nos gastos em consumo de Entretenimento e cultura (X_7).

Gráfico 4 – Representação da trajectória de Portugal

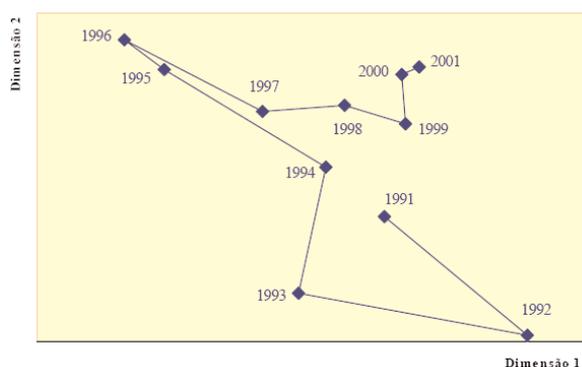


No que se refere à Dinamarca (Gráfico 5), o ano de 1991 é caracterizado por um crescimento nos gastos



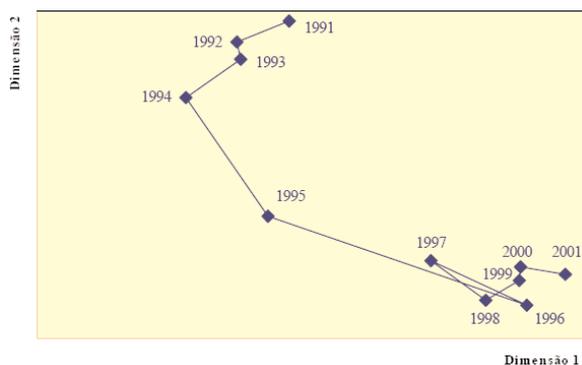
em consumo das famílias em Comunicações (X_6), ao que se segue um decréscimo nos gastos em consumo nesta rubrica até 1996. Nos anos de 1997 a 2001, os gastos das famílias dinamarquesas são caracterizados por um decréscimo em Vestuário e calçado (X_4) e por um aumento em Habitação, água, electricidade, gás e outros combustíveis (X_5).

Gráfico 5 – Representação da trajectória da Dinamarca



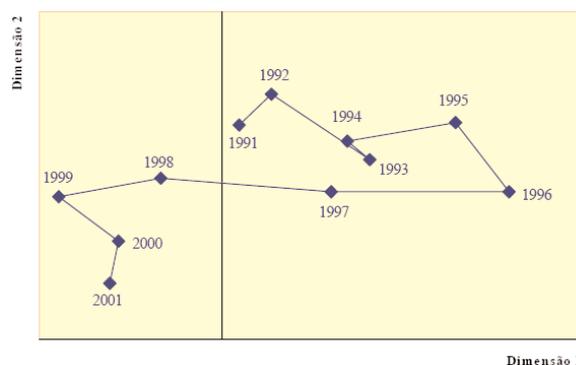
Ao longo de todo o período em análise e representado no gráfico 6, a evolução dos gastos em consumo das famílias holandesas caracteriza-se sobretudo pelo crescimento em Comunicações (X_6) e pelo decréscimo em Alimentação e bebidas não alcoólicas (X_1), Bebidas alcoólicas, tabaco e narcóticos (X_2) e Vestuário e calçado (X_4).

Gráfico 6 – Representação da trajectória da Holanda



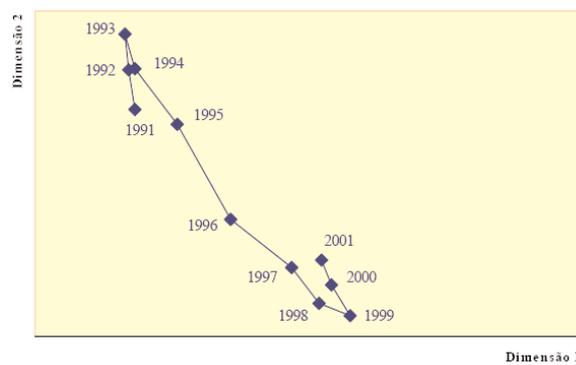
A observação do gráfico 7 leva à constatação de que no período compreendido entre 1991 e 1995, os gastos em consumo das famílias irlandesas em percentagem dos gastos totais em consumo em Entretenimento e cultura (X_7) decresceram. Por outro lado, cresceram de forma forte e *quasi*-linear, os gastos em consumo destas famílias em Habitação, água, electricidade, gás e outros combustíveis (X_5) no período 1996-1999 e em Comunicações (X_6) no período 1999-2001.

Gráfico 7 – Representação da trajectória da Irlanda



Na Finlândia (Gráfico 8), os gastos das famílias em Alimentação e bebidas não alcoólicas (X_1) são caracterizados por uma estagnação até ao ano 1993, ao que se seguiu um decréscimo até 1999. No período 1993-1999, observa-se também que os gastos em consumo cresceram nas rubricas Entretenimento e cultura (X_7) e Comunicações (X_6), e decresceram na componente Bebidas alcoólicas, tabaco e narcóticos (X_2). Os últimos 3 anos caracterizam-se por um crescimento dos gastos em Bebidas alcoólicas, tabaco e narcóticos (X_2) e por uma evolução constante nos gastos em Alimentação e bebidas não alcoólicas (X_1) e Entretenimento e cultura (X_7).

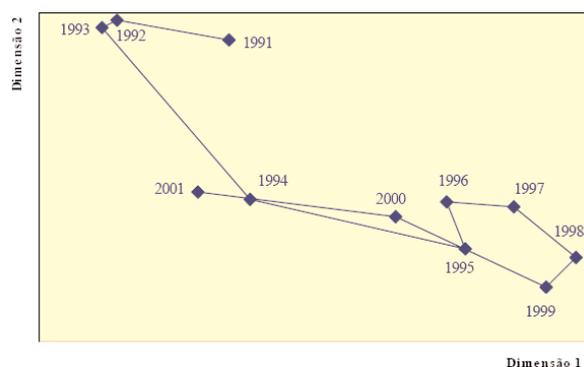
Gráfico 8 – Representação da trajectória da Finlândia



O gráfico 9 mostra que os gastos em consumo das famílias italianas em Habitação, água, electricidade, gás e outros combustíveis (X_5) cresceram até 1996, momento a partir do qual começaram a decrescer. A partir de 1993, observou-se ainda que os gastos destas famílias em Alimentação e bebidas não alcoólicas (X_1) decresceram, enquanto os gastos em Comunicações (X_6) aumentaram.

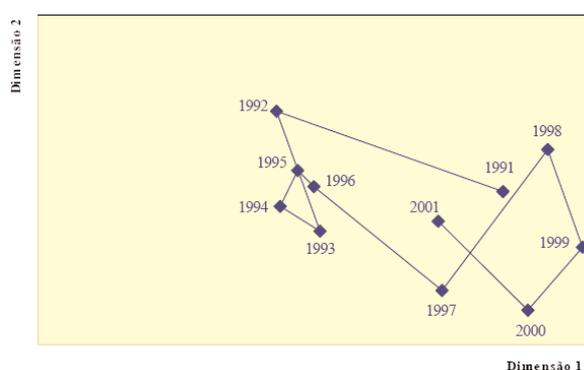


Gráfico 9 – Representação da trajectória da Itália



Os gastos em consumo das famílias austríacas são caracterizados por uma grande irregularidade ao longo de todo o período, sobretudo nas rubricas Bebidas alcoólicas, tabaco e narcóticos (X_2) e Entretenimento e cultura (X_7) (Gráfico 10).

Gráfico 10 – Representação da trajectória da Áustria



4. CONCLUSÃO

A DACP é uma técnica de análise de dados com grande potencial de aplicação a dados reais. A metodologia apresentada sugere diferentes perspectivas de análise que, em conjunto, permitem uma análise de dados muito rica.

Na realidade, a partir da utilização desta metodologia foi possível verificar a existência, no período em análise, de uma evolução temporal decrescente do peso dos gastos em consumo das famílias em alimentação e bebidas e em vestuário e calçado, relativamente aos gastos totais, isto é, verificou-se um decréscimo do peso dos bens de primeira necessidade, o que parece denotar uma melhoria do nível de vida das populações dos países em análise. Em oposição, nesse período registou-se uma subida dos gastos em habitação, água, electricidade, gás e outros combustíveis, comunicações e entretenimento e cultura. Também esta subida é indicadora da melhoria do nível das

populações, pelo menos no que diz respeito ao aumento dos gastos em comunicações (provavelmente em telemóveis e Internet), em entretenimento e cultura e com combustíveis.

Por outro lado, a análise por grupos de anos vem, de certa forma, confirmar o que anteriormente se afirmou. No entanto, entre 1995 e 1998, observou-se um decréscimo dos gastos em consumo das famílias em educação e em saúde, talvez como reflexo da crise económica que se viveu na Europa no início da década de 90. Por último, o crescimento nos gastos em consumo, em percentagem dos gastos totais em consumo, em alimentação e bebidas, educação e habitação, água, electricidade, gás e outros combustíveis e a diminuição dos gastos em entretenimento e cultura a partir do ano 2000 deverão ser objecto de análises futuras, conjuntamente com dados mais recentes, para se verificar se esta trajectória algo inexplicável dos gastos nessas variáveis se mantém.

No que diz respeito à análise por países, a oposição em termos de gastos em consumo observada entre os países do sul e os do norte da Europa parece fazer todo o sentido, especialmente se se tomar em conta que Portugal e Itália são os países que menos gastaram em entretenimento e cultura naquele período e os que mais gastaram em educação no período 1991-1995 e em alimentação e bebidas a partir de 1996. Por outro lado, os gastos dos finlandeses em entretenimento e cultura entre 1991 e 1995 foram os mais elevados e as suas despesas em alimentação, bebidas, vestuário e calçado foram as mais baixas. Acresce ainda que os dinamarqueses foram os que tiveram maiores despesas em entretenimento e cultura e menores em vestuário e calçado, entre 1991 e 2001.

No que diz respeito a Portugal, no período de preparação da economia para a entrada na UEM e que se caracterizou por uma convergência europeia, verificou-se uma diminuição dos gastos em consumo das famílias em alimentação, vestuário e calçado e uma evolução crescente nos gastos em entretenimento e cultura.

Por último, é importante realçar o facto da metodologia usada neste artigo, embora bastante útil na aplicação a problemas económicos, ser ainda pouco utilizada, o que torna este estudo pioneiro em Portugal. Mais ainda, o facto deste artigo ser um dos primeiros da literatura no qual são apre-



sentadas as ajudas clássicas à interpretação no contexto da DACP, constitui certamente uma mais valia. Em investigação futura, pretende-se aplicar esta metodologia a outros conjuntos de dados reais, com o objectivo de investigar a sua sensibilidade na análise de outras situações económicas.

5. BIBLIOGRAFIA

AMBAPOUR, S. (2001): *STATIS: Une methode d'analyse conjointe de plusieurs tableaux de donnes*, Document de travail (DT 01/2001), Bureau d'Application des Methodes Statistiques et Informatiques.

BOUROCHE, J.-M. (1975): *Analyse des données ternaires: la double analyse en composantes principales*, Thèse de 3^o cycle, Université de Paris VI, Paris.

BOUROCHE, J.-M. e DUSSAIX, A.-M. (1975): "Several alternatives for three-way data analysis", in: *Metra*, 14, pp. 299-319.

BOUROCHE, J.-M. e SAPORTA, G. (2002): *L'Analyse des Données*, Press Universitaires de France, Paris.

ESCOUFIER, Y. (1973): "Le traitement des variables vectorielles", in: *Biometrics*, 29, pp. 750-760.

EUROSTAT (s.d.): *Living conditions and social protection*, Luxembourg, Office for Official Publication of the European Communities [disponível em http://epp.eurostat.ec.europa.eu/portal/page?_pageid=3134,70318806,3134_70394055&_dad=portal&_schema=PORTAL, acedido 20/02/2007].

JOHNSON, R. A. e WICHERN, D. W. (2002): *Applied Multivariate Statistical Analysis* (5th ed.), Prentice-Hall, New Jersey.

KROONENBERG, P.M. (1983): *Three-mode principal component analysis*, DSWO Press, Leiden.

LAVIT, C. (1988): *Analyse conjointe de tableaux quantitatifs*, Masson, Paris.

LEBART, L., MORINEAU, A. e PIRON, M. (2000): *Statistique Exploratoire Multidimensionnelle* (3e éd.), Dunod, Paris.

SAPORTA, G. (1990): *Probabilités: Analyse des Données et Statistique*, Éditions Technip, Paris.

SAPORTA, G. (1996): *Analyse des données évolutives: Methodes et Applications*, Éditions Technip, Paris.

SILVA, C. (1997), Portugal e Moeda Única, Editorial Verbo, Lisboa.