

Descripció i classificació de les comarques catalanes en regions homogènies segons l'ús de la terra.

*Tomàs Aluja**

Résumé / Abstract / Resumen

Cet article se réfère à l'utilisation des techniques d'analyse des données de recherche nécessaires à l'étude des grands tableaux numériques qui correspondent à des données de type spatial. Fréquemment, face à un grand volume de données sur un territoire, par exemple dans l'élaboration d'un recensement de population, il est difficile d'assimiler toute l'information obtenue. Nous allons montrer comment ces techniques permettent une connaissance plus approfondie des données utilisées, sans référence directe aux tableaux de données.

Les objectifs fixés sont normalement: (1) mettre en relief les traits les plus caractéristiques des données, comme par exemple, associations et/ou oppositions des objets étudiés, ce qui est facilement obtenu à l'aide de méthodes d'analyse factorielle descriptive; (2) faire une classification des objets étudiés en peu de classes représentatives, ce qui est facile à obtenir en appliquant un algorithme rapide de classification ascendante hiérarchique.

L'application des techniques en question nous montre la compatibilité des deux types de résultats. Normalement, celle-là est la dernière étape de l'étude d'une table de données, en géographie et aussi dans d'autres sciences où les observations correspondent à des petites surfaces dans l'espace, comme la géologie, l'écologie... Il est normal d'essayer de faire correspondre les classes obtenues avec des régions géographiques. Pour cette raison, il a fallu introduire une relation de continuité dans l'algorithme de classification

* Departament d'Investigació Operativa i Estadística, Universitat Politècnica de Barcelona.

ascendante hierarchique. Leur application nous permet de voir comment peut s'améliorer l'interpretation des classes obtenues.

. . .

The theme of this article is the application of techniques of exploratory statistics to the study of comprehensive numerical tables consisting of statistics of a spatial nature. The immensity of statistics compiled over a large area, as in the case of a population census, frequently makes it difficult to assimilate all the information contained therein. It is shown that the mentioned techniques of analysis make possible a profound understanding of such statistics without resorting to the inspection of the said tables.

The objectives usually pursued are: (1) to emphasize the most outstanding characteristics of the statistics, such as associations and/or contrasts in the elements under study, an objective which is easily fulfilled through methods of descriptive factorial analysis; (2) to group the basic elements of study into a limited number of representative classes, which can likewise be easily achieved through a simple algorithm of ascendent hierarchical classification. The application of this method demonstrates the compatibility of the two results. This normally corresponds to the final stage in the study of statistical tables, in which observations relate to small areas points. The natural desire to make the classes obtained coincide with geographical regions made necessary the introduction of the content relationship within the algorithm of ascendent hierarchical classification. The application undertaken makes it possible to identify improvements in the interpretation of the classes obtained.

. . .

Este artículo trata de la utilización de las técnicas de análisis de los datos exploratorios para el estudio de grandes tablas numéricas, correspondientes a datos de tipo espacial. Muy a menudo la gran cantidad de datos recogidos sobre un territorio, por ejemplo en la confección de un Padrón de habitantes, hace difícil asimilar toda la información que contienen. Veremos como las técnicas mencionadas, permiten un conocimiento profundo de los datos puestos en juego sin necesidad de ninguna inspección de la tabla de los mismos.

Los objetivos que normalmente se persiguen son: Poner de manifiesto las características más destacadas de los datos, tales como asociaciones y/o oposiciones de los objetos bajo estudio, que es obtenido cómodamente mediante los métodos del análisis descriptivo factorial. Efectuar una clasificación de los objetos base del estudio, en unas pocas clases representativas, que así mismo también es obtenido fácilmente haciendo uso de un rápido algoritmo de clasificación jerárquica ascendente. La aplicación del método, muestra la compatibilidad de ambos resultados. Normalmente esta es la última fase del estudio de una tabla donde las observaciones corresponden a puntos o pequeñas superficies en el espacio, como Geología, Ecología... es normal querer hacer corresponder las clases obtenidas con regiones geográficas, por lo que ha sido necesario introducir la relación de contenido dentro del algoritmo de clasificación jerárquica ascendente. La aplicación efectuada permite ver en qué forma mejora la interpretación de las clases obtenidas.

Ens proposem mostrar a partir d'un senzill exemple, com les tècniques d'anàlisi de dades permeten d'assolir un coneixement profund del territori, en el cas de dades geogràfiques. Aquestes tècniques consisteixen en la utilització conjunta d'una anàlisi factorial descriptiva, seguida d'una classificació sobre els individus, que en el nostre cas representaran diferents unitats territorials. Normalment aquesta és l'última fase en un estudi del territori, nosaltres però, hem prosseguit amb l'objectiu de definir sobre el territori regions geogràfiques homogènies. Aquest pot ésser un objectiu per ell mateix o una fase prèvia d'un estudi porterior. L'article ha estat dividit en tres parts, corresponents cadascuna a les tres tècniques utilitzades:

1 - Anàlisi factorial descriptiva

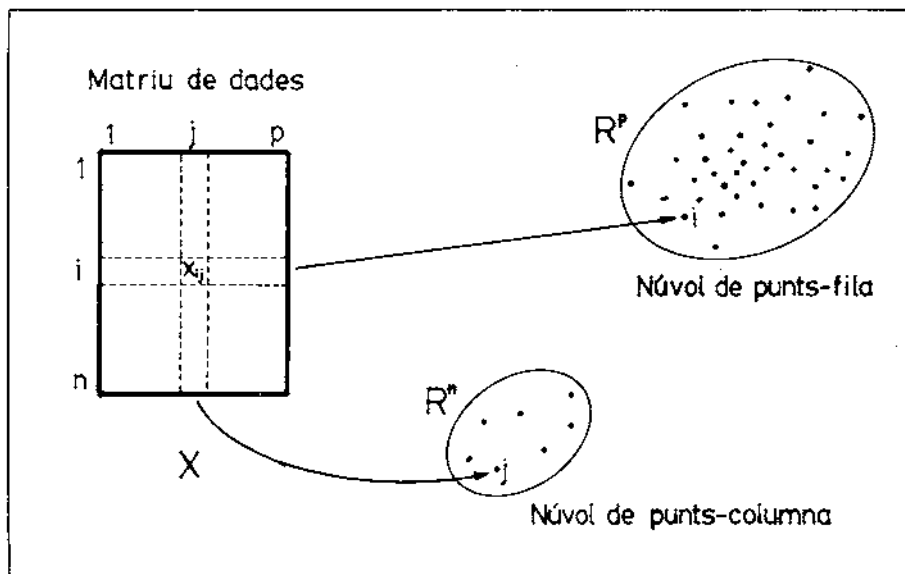
Sovint les dades recollides adopten la forma de grans taules numèriques, on les files corresponen a les unitats territorials estudiades i les columnes a les diferents característiques o variables mesurades per cada unitat territorial.

Cada fila de la taula la podem suposar com un punt de l'espai R^p de coordenades on les mesures de les variables: $(x_{ij}), j=1 \dots p$, constitueixen el conjunt de totes les files el núvol de punts-fila.

Anàlogament, podem suposar cada columna com un punt de l'espai R^n de coordenades i juntament amb els diferents valors mesurats per a totes les unitats territorials: $(x_{ij}), i=1 \dots n$, constitueixen el núvol de punts-columna.

Vegeu el gràfic 1, on x_{ij} és la mesura de la variable j en la unitat i .

Gràfic 1. Núvol de punts-fila en R^p i núvol de punts-columna en R^n



La interpretació de les proximitats entre punts en aquests dos núvols, és clara: dos punts-fila propers indicaran que es tracta de dues unitats territorials de característiques semblants, anàlogament dos punts-columna propers indicaran que aquelles dues variables tenen mesures semblants en totes les unitats territorials, això és, mesuren el mateix concepte. El problema rau en el caràcter n-dimensional dels núvols de punts, que els fan inassequibles a la comprensió humana. Les tècniques d'anàlisi factorial descriptiva tracten simplement de projectar el núvol de punts sobre plans factorials «privilegiats» resumint el màxim d'informació del núvol original. Aquest serà l'objecte de la primera part de l'article.

2 - Classificació automàtica

Les representacions planes obtingudes amb les tècniques precedents ja són útils, puix contenen amb tota seguretat els trets més característics de la taula de dades sota estudi. Sovint, però, l'investigador requereix una sintetització de la realitat en unes poques classes representatives del total, i així poder caracteritzar cada unitat territorial a partir de la seva classe de pertinença, aquest serà l'objectiu de la segona part de l'article.

3 - Classificació amb constricció de contigüitat

Les classes obtingudes pels algorismes de classificació automàtica precedents, són coherents, en el sentit d'agrupar unitats territorials homogènies entre si. En Geografia però i en altres ciències, és freqüent voler fer correspondre les classes obtingudes amb regions sobre el mapa, per la qual cosa cal introduir la restricció de contigüitat dins l'algorisme de classificació. Aquest ha estat l'objecte de la tercera part.

Les dades emprades

La taula de dades que ens proposem analitzar és extreta de l'estudi «L'agricultura catalana, estudi econòmic», donant per cadascuna de les 38 comarques, la utilització del sòl segons els 7 grans epígrafs següents mesurats en hectàries (vegeu la taula I):

- Superfície no agrícola	NOAG
- Superfície d'herbació de regadiu	HERE
- Superfície de llenyosos de regadiu	LERE
- Superfície d'herbació de secà	HESE
- Superfície de llenyosos de secà	LESE
- Prats i pastures	PAST
- Terreny improductiu i erm	IMPR

Aquests epígrafs resulten de l'agregació de varis tipus de conreus, (en majúscules es dona l'identificador de cada epígraf), en la taula es dona a més l'identificador i el número seqüencial corresponent a cada comarca.

Les reduïdes dimensions d'aquesta matriu (38*7), no justifiquen la utilització de

potents eines d'exploració, per la qual cosa seria suficient una acurada inspecció visual. No obstant això anem a procedir com si d'un problema de grandària real es tractés, pel qual en el nostre cas tindrem un núvol de 38 punts en R^7 corresponent a les comarques, i un núvol de 7 punts en R^{38} corresponent als diferents usos de la terra.

TAULA I. MATRIU DE DADES

		NOAG	HERE	LERE	HESE	LESE	PAST	IMPR	TOTAL
Baix Llobregat	1 BALL	14.093	3.172	2.131	1.972	5.826	8.170	12.113	47.477
Barcelonès	2 BARC	14.057	664	0	22	120	187	873	15.923
Maresme	3 MARE	8.118	4.996	174	3.392	1.815	4.125	16.650	39.270
Vallès Occidental	4 VAOC	16.022	1.907	90	7.466	2.507	3.919	30.796	62.707
Vallès Oriental	5 VAOR	11.852	4.964	105	10.363	1.262	14.656	41.786	84.988
Alt Empordà	6 AEMP	20.456	7.637	1.518	24.819	8.793	13.237	57.438	133.898
Baix Empordà	7 BEMP	10.353	6.571	1.061	19.189	1.425	1.193	30.256	70.048
Garrotxa	8 GART	6.673	61	23	9.234	67	5.116	52.246	73.420
Gironès	9 GIRO	6.782	2.965	500	30.753	522	2.065	45.939	89.526
La Selva	10 SELV	13.102	2.402	510	8.796	1.039	8.573	61.443	95.865
Alt Camp	11 ACAM	6.626	670	2.558	4.844	19.745	644	19.738	54.825
Alt Penedès	12 APEN	5.276	469	287	8.790	17.623	4.767	17.927	55.139
Baix Penedès	13 BPEN	5.020	268	55	2.522	7.654	18	10.869	26.406
Garraf	14 GARF	2.539	153	46	2.529	3.298	2.233	15.351	26.149
Tarragonès	15 TARR	8.048	634	2.917	645	12.262	176	9.820	34.502
Baix Camp	16 BCAM	8.494	3.721	5.799	1.688	20.482	900	26.712	67.796
Conca de Barberà	17 CBAR	4.449	540	83	17.165	19.608	1966	28.985	72.796
Priorat	18 PRIO	7.035	365	402	1.288	15.752	1.976	24.913	51.731
Ribera d'Ebre	19 RIBE	12.799	1.252	3.074	2.358	24.297	975	37.725	82.480
Baix Ebre	20 BEBR	13.137	16.400	2.432	466	38.904	3.227	31.165	105.731
Montsià	21 MONT	8.192	9.788	2.065	988	26.817	935	17.477	66.262
Terra Alta	22 TEAL	7.177	308	63	3.691	26.764	2.596	33.405	74.004
Cerdanya	23 CERD	12.126	2.128	32	6.031	1.849	15.600	22.488	60.254
Osona	24 OSON	15.212	568	12	26.540	10	15.613	60.615	118.570
Ripollès	25 RIPO	10.020	15	8	4.800	4	23.787	59.141	97.775
Anoia	26 ANOI	10.941	561	47	23.732	6.756	11.080	33.839	86.956
Bages	27 BAGE	15.407	1.488	29	18.666	2.697	25.593	61.492	125.822
Berguedà	28 BERG	16.980	160	6	8.769	16	34.714	52.052	112.697
Solsonès	29 SOLS	9.698	5	0	18.297	10	30.460	41.394	99.864
Les Garrigues	30 GARI	3.848	9.736	2.176	10.709	34.480	2.102	17.328	80.379
La Noguera	31 NOGU	23.293	24.105	3.717	49.554	7.810	16.059	56.756	181.294
Segarra	32 SEGA	4.794	416	580	44.217	3.767	5.325	14.165	73.264
Segrià	33 SEGR	12.779	44.742	27.634	25.487	16.589	4.911	14.765	146.907
Urgell	34 URGE	4.751	17.854	1.495	32.787	10.455	2.163	8.711	78.216
Alt Urgell	35 AURG	30.641	972	222	5.288	84	34.402	64.354	135.963
Pallars Jussà	36 PAJU	30.273	3.840	134	19.922	2.726	51.779	63.205	171.879
Pallars Sobirà	37 PASO	29.257	429	21	5.890	148	34.494	66.817	137.056
Vall d'Aran	38 VAAR	17.539	5	2	581	0	13.152	30.769	62.048
	TOTAL	457.859	176.931	62.008	464.250	343.983	402.888	1.291.968	3.199.887

Anàlisi factorial descriptiva: l'anàlisi de correspondències

Tractant-se d'una taula homogènia (totes les caselles vénen mesurades en la mateixa unitat), el mètode d'anàlisi factorial més adequat és l'anàlisi de correspondències. Els avantatges d'aquest mètode sobre el més conegut d'anàlisi en components principals deriven de les seves propietats, les quals podem resumir molt breument en les següents:

— Pondera cada comarca donant-els-hi una importància proporcional a la seva superfície total, (superfície d'una comarca $\sum_j x_{ij}$. Pes d'un punt-comarca $\frac{\sum_j x_{ij}}{\sum_i \sum_j x_{ij}}$).

— Les comarques són comparades segons llurs perfils, és a dir, a partir del percentatge de cada tipus d'utilització del sòl respecte la superfície total de la comarca, $(\frac{x_{ij}}{\sum_j x_{ij}}, j = 1 \dots p)$, i no a partir de les dades brutes (x_{ij}) . A la taula II es donen

els perfils i la ponderació per cada comarca.

De la mateixa manera podríem parlar de la comparació dels diferents usos de la terra, la qual és feta a partir dels perfils de les columnes $(\frac{x_{ij}}{\sum_i x_{ij}}, i = 1, \dots, n)$.

— La distància entre punts és la de Khi-quadrat i no l'euclidiana clàssica, la qual té com a conseqüència donar més importància, per exemple, a una diferència de 0.18 en el percentatge de «llenyosos de regadiu» entre la Segarra i el Segrià, que a una diferència de 0.18 en el percentatge de «terreny improductiu i erm» entre el Garraf i la Conca de Barberà; car el «terreny improductiu» és l'ús de la terra més abundant (el 40% de Catalunya), i per tant és normal que les xifres siguin altes i lògicament les diferències també poden ser-ho. Per contra els «llenyosos de regadiu» és la utilització del sòl més rara, i per tant els percentatges en totes les comarques són baixos, pel que si en algun cas ens trobem amb un valor alt (per exemple el Segrià), això és més significatiu que no pas la situació precedent.

— El paper totalment simètric de les files i les columnes, el qual té com efecte la representació simultània dels punts-fila i dels punts-columna, serà tractat més endavant.

Càlcul dels eixos factorials

L'objectiu de l'anàlisi de correspondències és de cercar en els núvols de punts, (definites prenent com coordenades dels punts llurs perfils i ponderats segons la seva importància) aquelles direccions que anomenarem d'ara endavant eixos factorials, que més informació resumeixen dels núvols de punts originals (concretament el que s'intenta maximitzar és la inèrcia projectada sobre cada eix, la qual pot interpretar-se com la informació continguda per cada eix, Benzecri-1980). Vegi's el gràfic II.

I a continuació projectar el núvol de punts sobre els eixos factorials trobats.

Donem a la taula III la llista dels valors propis corresponents als 6 eixos factorials

trobats, (cada valor propi és igual a la inèrcia projectada sobre cada eix, en conseqüència indica la informació resumida per cada eix factorial). Pot demostrar-se a més que tant el núvol de punts-fila com el núvol de punts columna estan continguts

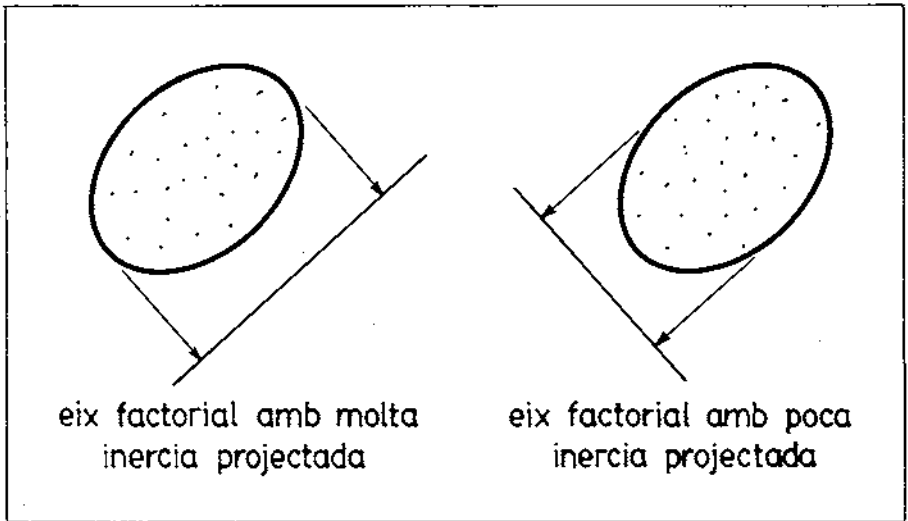
TAULA II. PERFILS DE LES COMARQUES

		NOAG	HERE	LERE	HESE	LESE	PAST	IMPRPES
BALL	0.297	0.067	0.045	0.041	0.123	0.172	0.255	0.0148
BARC	0.883	0.042	0.000	0.001	0.007	0.012	0.055	0.0050
MARE	0.207	0.127	0.004	0.086	0.046	0.105	0.424	0.0123
VAOC	0.255	0.030	0.001	0.119	0.040	0.062	0.491	0.0196
VAOR	0.139	0.058	0.001	0.122	0.015	0.172	0.492	0.0265
AEMP	0.153	0.057	0.011	0.185	0.066	0.099	0.429	0.0418
BEMP	0.148	0.094	0.015	0.274	0.020	0.017	0.432	0.0219
GART	0.090	0.001	0.000	0.126	0.001	0.070	0.712	0.0229
GIRO	0.076	0.083	0.005	0.343	0.006	0.023	0.513	0.0279
SELV	0.137	0.025	0.005	0.092	0.011	0.089	0.641	0.0299
ACAM	0.121	0.012	0.046	0.088	0.360	0.012	0.360	0.0171
APEN	0.096	0.008	0.005	0.159	0.320	0.086	0.325	0.0172
BPEN	0.190	0.010	0.002	0.095	0.290	0.001	0.412	0.0082
GARF	0.097	0.006	0.002	0.097	0.126	0.085	0.587	0.0082
TARR	0.233	0.018	0.084	0.019	0.355	0.005	0.285	0.0108
BCAM	0.125	0.055	0.085	0.025	0.302	0.013	0.394	0.0212
CBAR	0.061	0.007	0.001	0.236	0.269	0.027	0.398	0.0227
PRIO	0.136	0.007	0.007	0.025	0.304	0.038	0.481	0.0162
RIBE	0.155	0.015	0.037	0.028	0.294	0.012	0.457	0.0258
BEBR	0.124	0.155	0.023	0.004	0.368	0.030	0.295	0.0330
MONT	0.124	0.148	0.031	0.015	0.405	0.014	0.264	0.0207
TEAL	0.097	0.004	0.001	0.050	0.362	0.035	0.451	0.0231
CERD	0.201	0.035	0.001	0.100	0.031	0.259	0.373	0.0188
OSON	0.128	0.005	0.000	0.224	0.000	0.132	0.511	0.0370
RIPO	0.102	0.000	0.000	0.049	0.000	0.243	0.605	0.0305
ANOI	0.126	0.006	0.000	0.273	0.078	0.127	0.389	0.0272
BAGE	0.122	0.012	0.000	0.148	0.021	0.203	0.492	0.0393
BERG	0.150	0.001	0.000	0.078	0.000	0.308	0.462	0.0352
SOLS	0.097	0.000	0.000	0.183	0.000	0.305	0.414	0.0312
GARI	GARI	0.048	0.121	0.027	0.133	0.429	0.026	0.0251
NOGU	0.128	0.133	0.020	0.273	0.043	0.088	0.313	0.0566
SEGA	0.065	0.006	0.008	0.603	0.051	0.073	0.193	0.0229
SEGR	0.087	0.304	0.188	0.173	0.113	0.033	0.100	0.0459
URGE	0.061	0.228	0.019	0.419	0.133	0.028	0.111	0.0244
AURG	0.225	0.007	0.002	0.039	0.001	0.253	0.473	0.0425
PAJU	0.176	0.022	0.001	0.116	0.016	0.301	0.368	0.0537
PASO	0.213	0.003	0.000	0.043	0.001	0.252	0.487	0.0428
VAAR	0.283	0.000	0.000	0.009	0.000	0.212	0.496	0.0194
TOTAL	0.143	0.055	0.019	0.145	0.107	0.126	0.404	1.000

TAULA III. EDICIÓ DELS VALORS PROPIS

	VALOR PROPÍ	PERCENTATGE	PERCENTATGE	ACUMULAT
1	0.29325730	45.33	45.33	*****
2	0.16148239	24.96	70.30	*****
3	0.10106081	15.62	85.92	*****
4	0.03590611	5.55	91.47	****
5	0.03445056	5.33	96.80	****
6	0.02072324	3.20	100.00	**

en un mateix espai de dimensió 6, i que tenen idèntics valors propis. La suma de tots sis valors propis és la inèrcia total o equivalentment la informació continguda en el núvol original (de punts-fila o columna). El quocient entre la inèrcia de cada eix dividit per la inèrcia total ens dona el percentatge d'informació explicat per cada eix factorial; pel que en el nostre cas podem veure (vegeu la taula III) l'existència de tres direccions significatives, explicant el 45%, el 25% i el 16% respectivament de la inèrcia dels núvols originals.

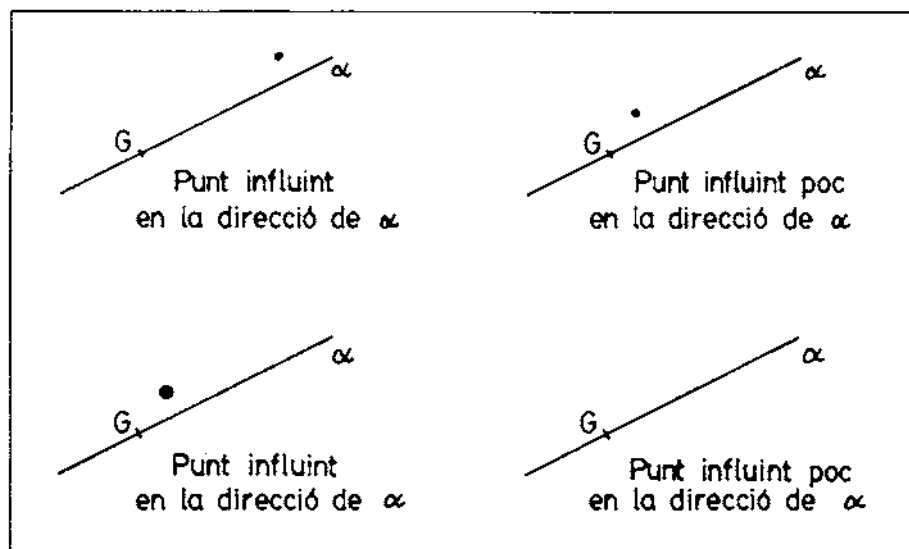


Gràfic II. Rectes de projecció

Interpretació dels eixos factorials

La direcció de cada eix factorial ve determinada per la posició dels punts (fila o columna) i per les seves ponderacions, vegi's per exemple el gràfic III.

Pel que serà interessant calcular la contribució de cada punt a la determinació de la direcció de cada eix factorial. Aquestes contribucions serveixen per interpretar els eixos trobats. En la taula IV donem el pes, la distància al quadrat al centre de gravetat en el núvol original, les coordenades en els tres primers eixos i les contribucions en la formació d'aquests tres primers eixos per cada punt-variable; atès que el normal és interpretar els eixos a partir de les variables i després explicar amb aquesta interpretació la posició de les comarques, per bé que podríem donar la mateixa informació corresponent als punts-fila. També es dona els cosinus quadrats de cada punt-variable amb els tres primers eixos factorials, els quals poden interpretar-se com la correlació (al quadrat) entre una variable i un eix factorial, obtinguda dividint el quadrat de la projecció de cada punt sobre un eix per la distància al quadrat d'aquest



Gràfic III. Punt influent i punt no influent

TAULA IV. COORDENADES I CONTRIBUCIONS DE LES VARIABLES

NOMS	PESOS	DISTAN- CIA	COORDENADES			CONTRIBUCIONS			COSINUS QUADRATS		
			F1	F2	F3	F1	F2	F3	F1	F2	F3
NOAG	0.143	0.28	-0.17	-0.11	-0.26	1.5	1.1	9.3	0.11	0.04	0.23
HERE	0.055	1.96	1.02	0.74	-0.45	19.6	18.8	10.9	0.53	0.28	0.10
LERE	0.019	4.50	1.48	0.75	-1.02	14.5	6.8	19.8	0.49	0.13	0.23
HESE	0.145	0.70	0.03	0.62	0.55	0.0	34.1	43.2	0.00	0.55	0.55
LESE	0.107	1.67	1.03	-0.72	0.23	39.0	34.8	5.4	0.64	0.31	0.03
PAST	0.126	0.66	-0.64	0.02	-0.30	17.6	0.0	10.9	0.62	0.00	0.13
IMPR	0.404	0.11	-0.24	-0.13	0.04	7.7	4.3	0.5	0.49	0.15	0.01

punt al centre de gravetat. És evident que un valor del cosinus quadrat alt (proper a la unitat), indica que aquell punt és proper a l'eix factorial, mentre que un valor del cosinus quadrat baix (proper a zero), indica que aquell punt és quasi-ortogonal a l'eix factorial. Així doncs, els cosinus quadrats donen la qualitat de representació dels punts sobre els eixos factorials i per tant, a l'hora de fer les interpretacions del gràfics obtinguts només serà lícit fer-ho a partir dels punts ben representats (amb valors dels cosinus quadrats alts).

Així podem dir que el 1er. eix factorial (F1) ve format en un 39% per la variable «lleyosos de secà», en un 19.66% per la variable «herbacis de secà», en un 17.6% per la variable «prats i pastures», en un 14.5% per la variable «lleyosos de regadiu», en un 7.7% per la variable «terreny improductiu i erm». A més ens adonem

que els punts LESE, HERE i LERE apareixen amb signe contrari, és a dir, oposats als punts PAST i IMPR, per la qual cosa és fàcil concloure que es tracta d'un eix oposant les comarques de conreus agraris, d'aquelles amb alts percentatges de superfícies de prats i de pastures i no productives.

El 2on. eix factorial es caracteritza per l'oposició de les variables (en parèntesi donem les seves contribucions a la formació del segon eix):

HESE (34.1%) contra LESE(34.8%)
 HERE (18.8%)
 LERE (18.8%)

És a dir, es tracta d'un eix agrari oposant les comarques segons siguin primordialment de conreus «llenyosos de secà», respecte tots els altres tipus de conreus, «herbacis de secà i de regadiu» i «llenyosos de regadiu».

El 3er. factor es caracteritza per l'oposició entre les variables:

HESE (43.2%) contra LERE (19.8%)
 HERE (10.9%)
 PAST (10.9%)
 NOAG (9.5%)

És a dir, aquest tercer factor separa les comarques d'«herbacis de secà» respecte les de «llenyosos de regadiu» i «herbacis de regadiu», que en el segon factor havien aparegut juntes.

Les representacions factorials obtingudes

Una propietat important de l'anàlisi de correspondència és la representació simultània dels punts-fila i dels punts columna: en efecte, pot demostrar-se que la posició d'un punt fila (=comarca en el nostre cas), és el baricentre de la posició de tots els punts-columna, ponderats pels percentatges de cada utilització del sòl en aquella comarca, multiplicat per una certa constant. $\psi_{i\alpha}$ coordenada de la comarca i en l'eix α , $\varphi_{j\alpha}$ coordenada de la variable j en l'eix factorial α)

$$\psi_{i\alpha} = K \frac{\sum_j x_{ij}}{\sum_j x_{ij}} \varphi_{j\alpha}$$

i a l'inrevés, la coordenada d'un punt-columna en un eix és el baricentre de les coordenades de tots els punts-fila, ponderades pel perfil de la columna i multiplicat per una constant.

$$\Psi_{j\alpha} = k \frac{\sum_i x_{ij}}{\sum_i x_{ij}} \Psi_{\alpha}$$

Aquestes fórmules permeten d'interpretar les proximitats entre punts corresponents als dos núvols de punts. Una comarca propera a un conjunt de variables indicarà que aquella comarca té un percentatge alt relativament parlant, de superfícies definides per aquelles variables. Anàlogament una utilització del sòl propera a un conjunt de comarques, indicarà que aquella utilització té percentatges elevats en el conjunt de comarques que l'envolten.

Primer plà factorial

Així doncs, en el 1er. pla factorial (vegeu el gràfic IV), (pla format pels dos primers eixos factorials), el qual resumeix el 70.3% de la inèrcia de tot el núvol (és per tant el pla més descriptiu), hi veiem que:

A l'extrem de la dreta hi apareixen una sèrie de comarques, Solsonès, Pallars Jussà, Berguedà.. al voltant de la variable PAST, la qual indica que és en aquestes comarques on els prats i els terrenys de pastura hi són predominants.

Al centre hi apareixen comarques caracteritzades per tenir un índex alt de terreny erm i no agrícola, tals com el Barcelonès, el Vallès Occidental... i també aquelles amb una distribució de les terres propera a la mitjana de tot Catalunya com l'Alt Empordà, l'Anoia...

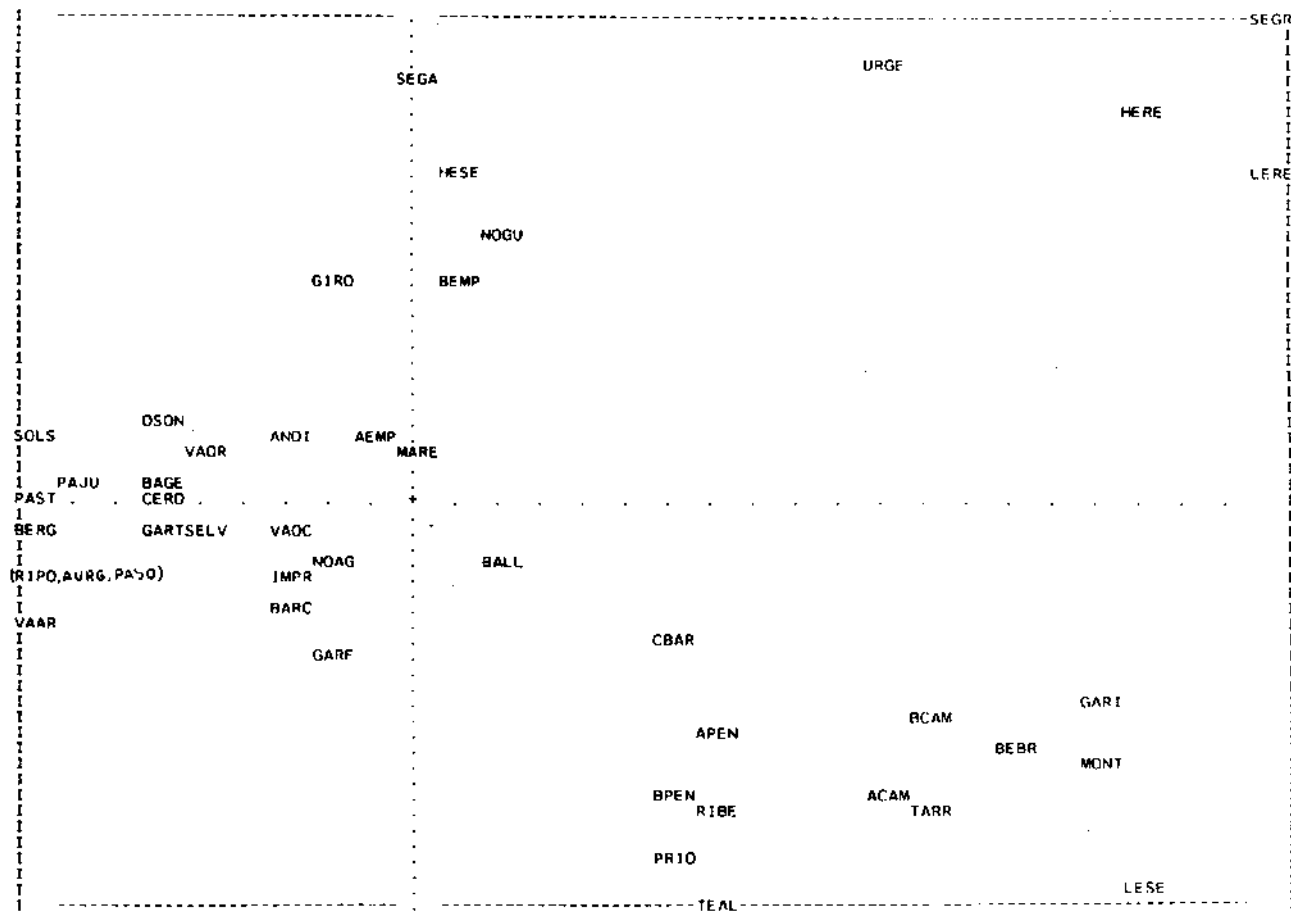
A l'extrem inferior dret hi apareixen tota una sèrie de comarques clarament separades, «estirades» per la variable LESE, la qual indica que és aquest tipus de conreu gairebé exclusiu d'aquestes comarques, totes elles pertanyents al sud de Catalunya, Terra Alta, Priorat, Tarragonès, Les Garrigues...

A l'extrem superior dret és la comarca del Segrià la que «atreu» als conreus de tipus herbacis de regadiu i llenyosos de regadiu, la qual cosa ens fa pensar que aquests conreus són gairebé exclusius d'aquesta comarca, més allunyada hi apareix la comarca de l'Urgell...

Segon pla factorial

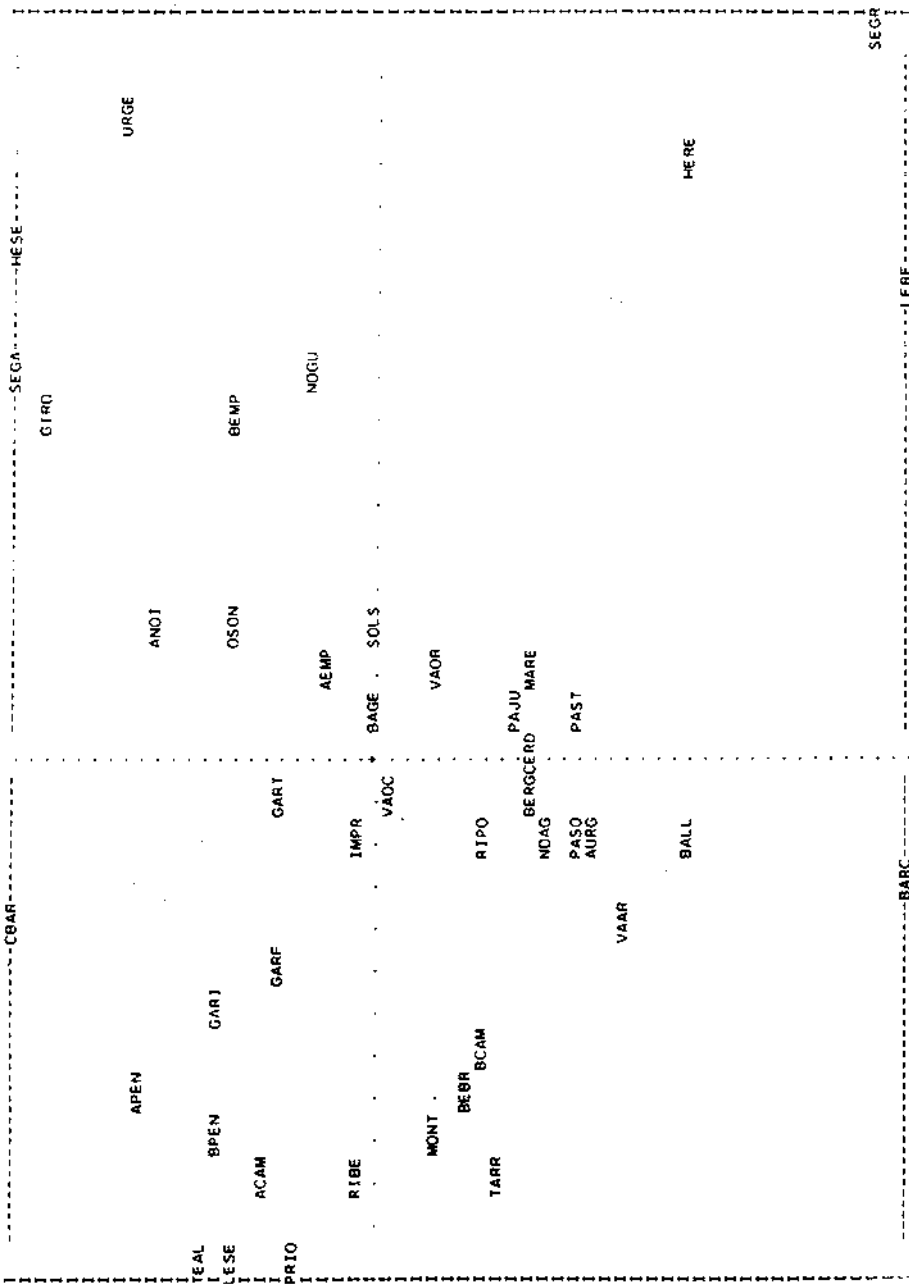
El segon pla factorial (vegeu el gràfic V), format pels eixos factorials 2on. i 3er, explica el 40.58% de la informació del núvol de punts original i separa el Segrià, que continua apareixent juntament amb les variables HERE i LESE, respecte les comarques de l'Urgell i la Segarra, que semblen ser les principals responsables de la posició del punt-variable HESE.

Les representacions factorials obtingudes representen bastant fidedignament la realitat, però des d'un punt de vista pràctic, a vegades és més útil disposar d'una agrupació de comarques en classes que siguin homogènies.



Gràfic IV. Representació simultània en el primer pla factorial

Gràfic V. Representació simultània en el segon pla factorial



CLASSIFICACIÓ AUTOMÀTICA: LA CLASSIFICACIÓ ASCENDENT JERÀRQUICA

Obtenció de l'arbre jeràrquic

Hem emprat el mètode de classificació ascendent jeràrquica (CAJ), segons el criteri de la variància, el qual és, per la riquesa d'informació que obté, el mètode més utilitzat. El funcionament de l'algorisme de la CAJ és iteratiu, comença per agregar en un sol element (que anomenarem nus), les dues comarques més semblants segons la utilització del sòl, en el nostre cas el Pallars Sobirà i l'Alt Urgell, formant el nus Núm. 39, i éssent l'índex de nivell la mesura de la dissemblança de les dues comarques agregades. En la següent iteració es procedeix anàlogament amb la diferència que només resten 37 punts per agregar, en el nostre cas els punts (=comarques) més propers resulten ser el Baix Ebre i el Montsià, formant un nus 40, etc... i així fins arribar a agregar totes les comarques en un sol nus Núm. 75. En el gràfic VI donem l'arbre jeràrquic bastit, en el qual es pot seguir tot el procés d'agregació. Observem que pel procés d'agregació seguit els índexs de nivell només poden ésser creixents.

Partició de l'arbre jeràrquic

L'arbre jeràrquic, també anomenat dendrograma, té doncs, interès per si mateix, però en la pràctica s'utilitza com un simple intermediari per determinar el nombre de classes, això és, agrupacions de comarques similars existents en les dades: per la qual cosa cal «tallar» l'arbre on existeixin salts importants de l'índex de nivell, (recordem que un índex de nivell alt indica que s'agreguen dos elements —comarques o nusos— dissemblants), per la qual cosa si s'ha produït un salt en el valor de l'índex de nivell, vol dir que s'ha passat d'agregar elements relativament similars a agregar-ne de dissimilars. Per la qual cosa és útil servir-se de l'histograma dels índexs de nivell, que donem la taula V, el qual permet detectar ràpidament els creixements sobtats de l'índex de nivell.

En el nostre cas això succeeix en agregar els nusos 74 i 73, 69 i 72, 68 i 70...

És segur que la participació de Catalunya en dues grans classes serà clarament interpretable i tindrà un sentit, però d'allò que es tracte és d'obtenir la partició més propera a la realitat, això és, amb el màxim nombre de classes possibles diferenciades i clarament interpretables. En conseqüència hem procedit a «tallar» l'arbre a l'altura del nus 72, obtenint les següents 5 classes:

CLASSE 1. Formada per les comarques del Baix Llobregat, Barcelonès, Maresme, Vallès Occidental, Alt Empordà, Baix Empordà, Gironès, Osona, Anoia i la Noguera.

CLASSE 2. Formada per les comarques del Vallès Oriental, la Garrotxa, La Selva, Garraf, Cerdanya, Ripollès, Bagès, Berguedà, Solsonès, Alt Urgell, Pallars Jussà, Pallars Sobirà i Vall d'Aran.

CLASSE 3. Formada per les comarques de l'Alt Camp, Alt Penedès, Baix Penedès, Tarragonès, Baix Camp, Conca de Barberà, Priorat, Ribera d'Ebre, Baix Ebre, Montsià, Terra Alta i Les Garrigues.

TAULA V. HISTOGRAMA DELS ÍNDEXS DE NIVELL DE LA CLASSIFICACIÓ JERÀRQUICA

NUS	ÍNDEX	ELEMENTS		HISTOGRAMA DELS ÍNDEXS DE NIVELL	
		AGREGATS	EFFECTIU		
39	0.0000	37	35	2	*
40	0.0003	20	21	2	*
41	0.0004	36	23	2	*
42	0.0004	18	13	2	*
43	0.0007	8	10	2	*
44	0.0007	42	19	3	*
45	0.0008	39	38	3	*
46	0.0009	27	5	2	*
47	0.0011	43	14	3	*
48	0.0011	17	12	2	*
49	0.0012	16	15	2	*
50	0.0013	41	28	3	*
51	0.0014	44	22	4	*
52	0.0015	51	11	5	*
53	0.0017	4	3	2	*
54	0.0017	24	26	2	*
55	0.0018	31	7	2	*
56	0.0020	6	53	3	*
57	0.0023	50	29	4	*
58	0.0028	25	46	3	*
59	0.0034	40	30	3	**
60	0.0040	9	54	3	**
61	0.0052	56	1	4	**
62	0.0061	49	52	7	***
63	0.0064	58	47	6	***
64	0.0078	57	45	7	***
65	0.0089	62	48	9	****
66	0.0112	55	61	6	****
67	0.0126	60	66	9	*****
68	0.0146	32	34	2	*****
69	0.0190	63	64	13	*****
70	0.0219	2	67	10	*****
71	0.0220	65	59	12	*****
72	0.0362	68	70	12	*****
73	0.0869	69	72	25	*****
74	0.1217	71	33	13	*****
75	0.2351	74	73	38	*****

CLASSE 4. Formada per les comarques de la Segarra i l'Urgell.

CLASSE 5. Formada per la comarca del Segrià.

Interpretació de la partició obtinguda

A continuació no resta sinó d'interpretar la partició obtinguda, operació aquesta sovint oblidada pels sistemes de classificació, i que nosaltres efectuem de forma automàtica tot comparant les mitjanes de cada variable en cada classe, respecte les mitjanes generals de les variables en tota la població, tenint en compte les respectives desviacions tipus.

Sabem que la variable:

$$t = \frac{x_k - x_j}{s_k}$$

on x_j és la mitjana general de la variable j . x_k és la mitjana de la variable j en la classe K . s_k és la desviació tipus de x_k segueix una distribució de Student, per tant, per interpretar una classe calcularem el valor de t per a cada varia-

TAULA VI. CARACTERITZACIÓ DE LA PARTICIPACIÓ EN 5. CLASSES

DESCRIPCIÓ DE LES CLASSES	t	PROVA	MITJANA		DESVIACIÓ TIPUS	
			CLASSE	GENERAL	CLASSE	GENERAL
CLASSE 1 (10 Comarques)						
HESE (Herbació de secà)	1.916	0.028	18743.8	12217.0	14869.1	
LESE (Llenyosos de secà)	-1.891	0.029	3558.4	9052.1	3212.0	10558.2
CLASSE 2 (13 Comarques)						
PAST (Prats i pastures)	4.248	0.000	22658.3	10602.2	13783.0	12448.9
IMP (Terreny improductiu)	3.379	0.000	48691.3	33999.1	16352.2	19070.7
LESE (Llenyosos de secà)	-3.339	0.000	1015.3	9052.1	1186.5	10558.2
NOAG (Superfície no agrícola)	2.362	0.009	15854.3	12048.8	8689.1	7068.1
HERE (Herbació de regadiu)	-1.739	0.041	1278.6	4656.0	1564.6	8518.9
LERE (Llenyosos de regadiu)	-1.514	0.065	87.5	1631.7	136.9	4474.3
CLASSE 3 (12 Comarques)						
LESE (Llenyosos de secà)	5.080	0.000	22032.3	9052.1	8488.2	10558.2
PAST (Prats i pastures)	-2.958	0.002	1690.1	10602.2	1318.6	12448.9
NOAG (Superfície no agrícola)	-2.655	0.004	7508.4	12048.8	2835.0	7068.1
HESE (Herbació de secà)	-2.543	0.006	4596.1	12217.0	4900.7	12386.0
IMPR (Terreny improductiu)	-2.382	0.009	23005.3	33999.1	8489.8	119070.7
CLASSE 4 (2 Comarques)						
HESE (Herbació de regadiu)	3.043	0.001	38502.0	12217.0	5715.0	12386.0
IMPR (Terreny improductiu)	-1.696	0.045	11438.0	33999.1	2727.0	19070.7
NOAG (Superfície no agrícola)	-1.476	0.070	4772.5	12048.8	21.5	7068.1
CLASSE 5 (1 Comarca)						
LERE (Llenyosos de regadiu)	5.811	0.000	27634.0	1631.7	0.0	4474.3
HERE (Herbació de regadiu)	4.705	0.000	44742.0	4656.0	0.0	8518.9

ble, a continuació ordenarem les variables segons els valors de t més extrems o de nivells de significació més petits. Les primeres variables de la llista són aleshores les més característiques d'aquella classe (les que més diferència hi ha entre la mitjana de la classe i la mitjana general, compte obtingut de les seves desviacions tipus).

Donem a la taula VI els resultats obtinguts per a les 5 classes detectades.

La qual cosa permet d'interpretar molt còmodament les cinc classes obtingudes:

La 1a. classe, formada per 10 comarques, es caracteritza per no tenir valors extrems en cap de les 7 variables considerades, si de cas només un percentatge lleugerament superior a la mitjana de tot Catalunya en «herbacis de secà» i una mica inferior en «llenyosos de secà». Aquestes comarques corresponen al grup de comarques que apareixen en una posició central en el 1er. pla factorial, les quals són les més industrials, o bé amb una utilització de la terra equilibrada entre els 7 usos considerats.

La segona classe, formada per 13 comarques, es caracteritza sobretot per tenir un índex alt de «prats i pastures» i «terreny improductiu», i per contra un percentatge baix en conreus «llenyosos de secà». Aquestes comarques corresponen al grup situat a l'esquerra del primer pla factorial, i són gairabé totes elles comarques de muntanya.

La tercera classe, formada per 12 comarques, es caracteritza especialment per tenir un percentatge alt de conreus «llenyosos de secà», mentre els percentatges de les altres utilitzacions són inferiors al promig de tot Catalunya. Aquestes comarques apareixen clarament separades en l'extrem inferior dret de el primer pla factorial, corresponen a les comarques agrícoles del sud de Catalunya, dedicades principalment a conreus de vinya, ametllers, olivers, avellaners...

La quarta classe és formada per dues comarques, la Segarra i l'Urgell, caracteritzades per tenir un percentatge elevat de conreus d'«herbacis de secà», tals com cereals i farratges. Aquestes comarques apareixen agrupades pel 3er. factor.

La quinta classe, formada només per la comarca del Segrià, es caracteritza per tenir uns percentatges molt alts de conreus de regadiu, tant llenyosos com herbacis. Això s'explica per la presència del canal d'Urgell, el qual determina la importància que tenen en aquesta comarca els cultius de regadiu, sobretot arbres fruiters, la qual cosa fa del Segrià una comarca a part des del punt de vista d'utilització de la terra, la qual cosa ve confirmada per la posició isolada d'aquesta comarca en els plans factorials mostrats.

Aquesta partició ha estat cartografiada en el gràfic VII adjunt.

CLASSIFICACIÓ AMB RESTRICCIÓ DE CONTINGÜITAT

La classificació obtinguda, si bé és coherent des del punt de vista de l'homogeneïtat de les classes obtingudes, és irreal si el que ens interessa és palesar les regions geogràfiques homogènies que puguin existir, ja que per exemple el Garraf apareix juntament amb comarques com el Solsonès, Berguedà... o bé La Noguera, la comarca més extensa de Catalunya, envoltada de comarques de diferents classes: de muntanya, ceralleres, el Segrià; participa de les característiques de totes elles... i surt agrupada amb la classe central, és a dir, amb el Baix Llobregat, Barcelonès...

Si existeix un interès en obtenir classes que corresponguin a regions en l'espai, cal introduir la relació de contigüitat (fenòmen extern) dins el procés d'agregació.

Nosaltres hem utilitzat l'algorisme de classificació ascendent jeràrquica del sistema SPAD, el qual segueix la metodologia presentada per Lebart (1978), segons la qual es tracta de limitar en cada iteració la recerca de la parella realitzant l'òptim dins les comarques contigües, essent la relació de contigüitat introduïda en codificació reduïda, això és, donant per cada comarca les adreces (números) de les seves comarques veïnes. A la taula VII donem la matriu de contigüitat de les comarques de Catalunya. Lògicament després de cada iteració s'efectua la posta al dia de la relació de contigüitat.

TAULA VII
Matriu de contigüitat de les comarques catalanes

Comarca	Adreces de les comarques contigües							
* BALL - 1 *	2	4	12	14	26	27		
* BARC - 2 *	1	3	4	5				
* MARE - 3 *	2	5	10					
* VAOC - 4 *	1	2	5	27				
* VAOR - 5 *	2	3	4	10	24	27		
* AEMP - 6 *	7	8	9					
* BEMP - 7 *	6	9	10					
* GART - 8 *	6	9	10	24	25			
* GIRO - 9 *	6	7	8	10				
* SELV - 10 *	3	5	7	8	9	24		
* ACAM - 11 *	12	13	15	16	17	26		
* APEN - 12 *	1	11	13	14	26			
* BPEN - 13 *	11	12	14	15				
* GARF - 14 *	1	12	13					
* TARR - 15 *	11	13	16					
* BCAM - 16 *	11	15	17	18	19	20		
* CBAR - 17 *	11	16	18	26	30	32	34	
* PRJO - 18 *	16	17	19	30				
* RIBE - 19 *	16	18	20	22	30	33		
* BEBR - 20 *	16	19	21	22				
* MONT - 21 *	20							
* TEAL - 22 *	19	20						
* CERD - 23 *	25	28	35					
* OSON - 24 *	5	8	10	25	27	28		
* RIPO - 25 *	8	23	24	28				
* ANOI - 26 *	1	11	12	17	27	29	32	
* BAGE - 27 *	1	4	5	24	26	28	29	32
* BERG - 28 *	23	24	25	27	29	35		
* SOLS - 29 *	26	27	28	31	32	35		
* GARI - 30 *	17	18	19	33	35			
* NOGU - 31 *	29	32	33	34	35	36		
* SEGA - 32 *	17	26	29	31	34			
* SEGR - 33 *	19	30	31	34				
* URGE - 34 *	17	30	31	32	33			
* AURG - 35 *	23	28	29	31	36	37		
* PAJU - 36 *	31	35	37	38				
* PASO - 37 *	35	36	38					
* VAAR - 38 *	36	37						

Com a observació val a dir que la codificació reduïda de la matriu de contigüitat permet importants estalvis en l'ocupació de memòria i en el temps de càlcul, puix solament cal limitar-se a investigar les parelles contigües.

A continuació donem l'histograma dels índexs de nivell amb el procés d'agregació seguit (vegeu la taula VIII): A la primera iteració s'agreguen igual que abans les comarques Núm. 35 i 37 corresponents a l'Alt Urgell i al Pallars Sobirà, atès que ambdues comarques són contigües, formant el nus 39. A la segona iteració també s'agreguen, com abans, el Baix Ebre i el Montsià (Núm. 20 i 21), formant el nus 40...

TAULA VIII. HISTOGRAMA DELS ÍNDEXOS DE NIVELL DE LA CLASSIFICACIÓ ASCENDENT JERÀRQUICA EN RESTICCIÓ DE CONTIGÜITAT

NUS	ÍNDEX	ELEMENTS		HISTOGRAMA DELS ÍNDEXOS DE NIVELL
		AGREGAT	EFFECTIU	
39	0.0000	35	37 2	*
40	0.0003	20	21 2	*
41	0.0006	18	19 2	*
42	0.0007	8	10 2	"
43	0.0008	38	39 3	*
44	0.0009	5	27 2	*
45	0.0010	12	13 2	"
46	0.0011	23	28 2	*
47	0.0012	15	16 2	*
48	0.0014	22	41 3	*
49	0.0016	11	45 3	*
50	0.0020	7	9 2	*
51	0.0021	29	46 3	*
52	0.023	24	44 3	*
53	0.0027	17	49 4	*
54	0.0034	25	51 4	**
55	0.0035	26	52 4	**
56	0.0035	14	53 5	**
57	0.0039	4	55 5	**
58	0.0037	3	57 6	**
59	0.0042	48	56 8	**
60	0.0047	36	43 4	**
61	0.0041	54	60 8	**
62	0.0047	6	50 3	**
63	0.0070	42	58 8	***
64	0.0090	47	59 10	****
65	0.0093	1	63 9	****
66	0.0094	31	34 2	****
67	0.0099	62	65 12	****
68	0.0108	30	64 11	****
69	0.0161	40	68 13	*****
70	0.0185	32	66 3	*****
71	0.0224	2	67 13	*****
72	0.0453	61	71 21	*****
73	0.0779	33	70 4	*****
74	0.1383	69	73 17	*****
75	0.2184	72	74 38	*****

En general podem observar que el procés d'agregació és força semblant a l'obtingut sense la restricció de contigüitat. És a dir, la restricció de contigüitat «força» poc el procés d'agregació, la qual cosa significa que cada comarca tendeix a tenir

característiques similars a les seves veïnes, això és, existeix un efecte geogràfic clar al poder establir una correspondència entre els punts de l'espai factorial i els punts del mapa geogràfic, també indica que i hauran inversions en els índex de nivell trobats, com és el cas del nus 61, (recordem que en la CAJ sense restricció els índexs de nivell sempre són creixents).

Partició de l'arbre obtingut

De l'observació dels índex de nivell (taula VIII), veiem que s'hi observa un fort creixement a partir del nus 72. Tallant, doncs, abans de formar aquest nus obtenim 5 classes clarament diferenciades. L'algorisme mateix ens dóna la seva composició a la taula IX.

TAULA IX

Partició en 5 classes de la classificació ascendent jeràrquica amb restricció de contigüitat

Tall de l'arbre jeràrquic pel nus 71

Classe	Elements que resten per agregar													
1	33	SEGR												
2	61	RIPO	SOLS	CERD	BERG	PAJU	VAAR	AURG	PASO					
3	69	BEBR	MONT	GARI	TARR	BCAM	TEAL	PRIO	RIBE	GARF	CBAR	ACAM	APEN	BPEN
4	70	SEGA NOGU URGE												
5	71	BARC	AEMP	BEMP	GIRO	BALL	GART	SELV	MARE	VAOC	OSON	VAOR	BAGE	

Això és:

CLASSE 1. Formada per la comarca del Segrià.

CLASSE 2. Formada per les comarques del Ripollès, Solsonès, Cerdanya, Berguedà, Pallars Jussà, Vall d'Aran, Alt Urgell i Pallars Sobirà.

CLASSE 3. Formada per les comarques del Baix Ebre, Montsià, les Garrigues, Tarragonès, Baix Camp, Terra Alta, Priorat, Ribera d'Ebre, Garraf, Conca de Barberà, Alt Camp, Alt Penedès i Baix Penedès.

CLASSE 4. Formada per les comarques de la Segarra, Noguera i Urgell.

CLASSE 5. Formada per les comarques del Barcelonès, Alt Empordà, Baix Empordà, Gironès, Baix Llobregat, Garrotxa, la Selva, Maresme, Vallès Occidental, Anoia, Osona, Vallès Oriental i Bages.

Interpretació de la partició obtinguda

Igual que abans, interessarà efectuar la interpretació de les cinc classes contigües obtingudes, la qual realitzem mitjançant l'estadístic t, obtenint els següents resultats presentats la taula X:

Dels valors obtinguts per l'estadístic t en aquesta taula, podem concloure que:

El Segrià continua mantenint-se com una classe a part, caracteritzada pels conreus de regadiu.

TAULA X. CARACTERITZACIÓ DE LA PARTICIÓ EN CINQ CLASSES
DE LA CLASSIFICACIÓ ASCENDENT JERÀRQUICA EN RESTICCIÓ DE CONTIGÜITAT

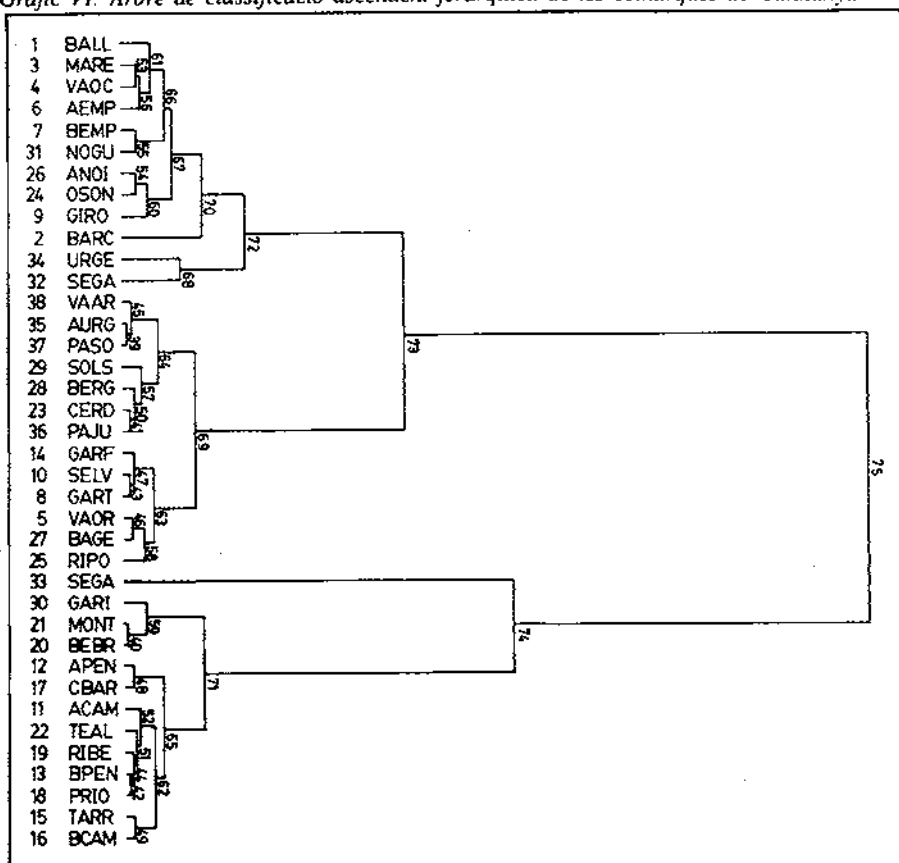
DESCRIPCIÓ DE LES CLASSES	I	PROVA	MITJANA		DESVIACIÓ TIPUS	
			CLASSE	GENERAL	CLASSE	GENERAL
CLASSE 1 (1 Comarca)						
LERE (Llenyosos de regadiu)	5.811	0.000	27634.0	1631.7	0.0	4474.3
HERE (Herbacis de regadiu)	4.705	0.000	44742.0	4656.0	0.0	8518.9
CLASSE 2 (8 Comarques)						
PAST (prats i pastures)	4.844	0.000	29798.4	10602.2	11542.4	12448.9
NOAG (Superfície no agrícola)	3.341	0.000	19566.7	12048.8	8557.6	7068.1
IMPR (Terreny improductiu)	2.640	0.004	50027.4	33999.1	15619.6	19070.7
LESE (Llenyosos de secà)	-2.513	0.006	604.6	9052.1	997.1	10558.2
HERE (Herbacis de regadiu)	-1.369	0.086	944.2	4656.0	1289.1	8518.9
CLASSE 3 (13 Comarques)						
LESE (Llenyosos de secà)	4.794	0.000	20591.2	9052.1	9561.8	10558.2
PAST (Prats i pastures)	-3.125	0.001	1731.9	10602.2	1275.1	12448.9
NOAG (Superfície no agrícola)	-3.055	0.001	7126.1	12048.8	3028.6	7068.1
HESE (Herbacis de secà)	-2.755	0.003	4437.1	12217.0	4740.6	12386.0
IMPR (Terreny improductiu)	-2.664	0.004	22416.5	33999.1	8407.9	19070.7
CLASSE 4 (3 Comarques)						
HESE (Herbacis de secà)	4.309	0.000	42186.0	12217.0	6994.1	12386.0
HERE (Herbacis de regadiu)	1.979	0.024	14125.0	4656.0	10024.0	8518.9
CLASSE 5 (13 Comarques)						
LESE (Llenyosos de secà)	-2.711	0.003	2526.0	9052.1	2714.5	10558.2

La segona classe, formada per 8 comarques, es caracteritza per tenir alts percentatges de «prats i pastures», «superfície no agrícola» i també «terreny improductiu». Aquesta classe equival a la segona trobada anteriorment, sense les comarques del Garraf, Bages, Vallès Oriental i Garrotxa, quedant limitada a les comarques pirenaïques exclusivament.

La tercera classe, formada per 13 comarques, es caracteritza per tenir alts percentatges de superfícies de «llenyosos de secà», i mancança de «prats i pastures», «superfície no agrícola»... Són les comarques agrícoles del sud de Catalunya detectades precedentment més el Garraf.

La quarta classe, formada per 3 comarques, es caracteritza pels conreus d'«herba-

Gràfic VI. Arbre de classificació ascendent jeràrquica de les comarques de Catalunya



cis de secà», és l'antiga classe formada per la Segarra i l'Urgell, a les quals s'ha agregat la Noguera.

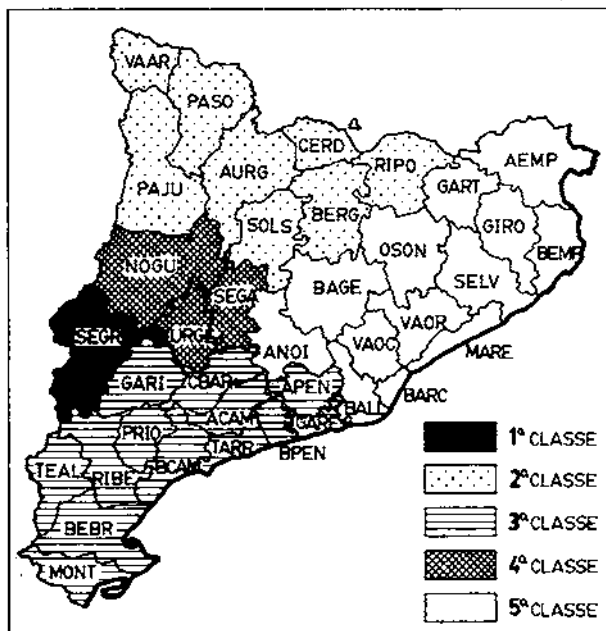
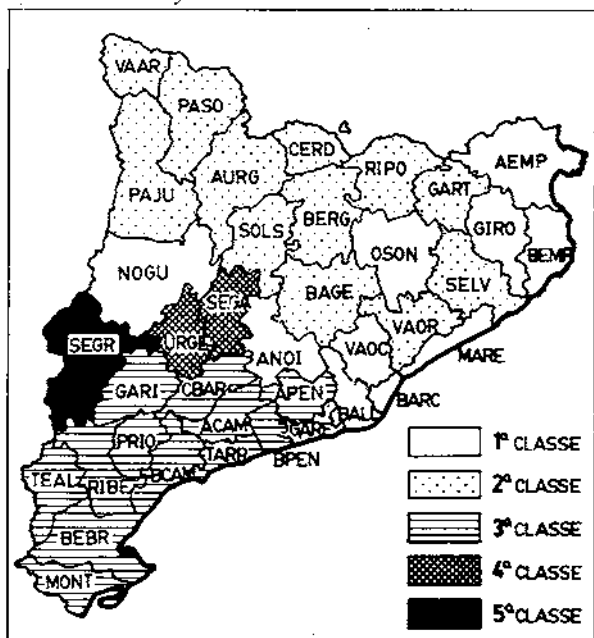
La cinquena classe, formada per 13 comarques, es caracteritza per no tenir valors extrems en cap dels diferents usos considerats, excepte en «llenyosos de secà» on el seu percentatge és inferior al total de Catalunya.

Com a conclusió global podem dir que es tornen a identificar «grosso modo» les mateixes classes d'abans, amb l'avantatge que ara les classes corresponen a regions geogràfiques, les quals hem cartografiat en el gràfic VIII.

Per altra banda la qualitat de la partició obtinguda no ha minvat, com ho proven els valors elevats de l'estadístic t, en algun cas superiors als obtinguts precedentment, sembla doncs clara l'existència a Catalunya de cinc regions diferenciades segons l'ús de la terra.

És clar, però, que un estudi amb unitats territorials més petites, com poden ésser els municipis, donaria amb més precisió les diferents regions agrícoles existents, car en una mateixa comarca s'hi poden donar característiques oposades.

Gràfic VII. Partició de Catalunya en cinc classes



Gràfic VIII. Partició de Catalunya en cinc classes amb restricció de contigüitat

Bibliografia

- ALUJA, T. 1984. *Mètodes de classificació i anàlisi factorial sobre un graf. Aplicació a l'anàlisi de dades municipals de Catalunya: Contribució a l'estudi de la divisió territorial*. Tesis doctoral. Setembre 1984. Universitat Politècnica de Catalunya.
- BENZECRI, J.P. 1973. *L'analyse des donnes. Tome 1: La Taxonomie, Tome 2: L'analyse des correspondances*. Dunod. París.
- BENZECRI, J.P. et coll. 1980 *Pratique de l'analyse des données. Tome 1: Analyse des correspondances, exposé élémentaire, Tome 2: Abrégé théorique*. Dunod. París.
- GRATACOS et al. 1980. *L'agricultura catalana, estudi econòmic*. Fundació CEP. Banca Catalana, 1980, Barcelona.
- LEBART, L., MORINEAU, A., TABARD, N. 1977. *Techniques de la description statistique*. Dunod. París.
- LEBART, L. 1978. «Programme d'agrégation avec contraintes». *Les cahiers de l'analyse des données*. Núm. 3, 1978, pp. 275-287.
- LEBART, L., MORINEAU, A., FENELON, J.P. 1985. *Tratamiento estadístico de datos. Métodos y programas*. Marcombo-Boixareu. Barcelona.
- SPAD. 1985. *Système portable pour l'analyse des données*. CESIA. 82 rue de Sèvres. 75007 París.