REVUE DE STATISTIQUE APPLIQUÉE

BRIGITTE ESCOFIER

Une représentation des variables dans l'analyse des correspondances multiples

Revue de statistique appliquée, tome 27, n° 4 (1979), p. 37-47

<http://www.numdam.org/item?id=RSA_1979__27_4_37_0>

UNE REPRÉSENTATION DES VARIABLES DANS L'ANALYSE DES CORRESPONDANCES MULTIPLES

Brigitte ESCOFIER*

RESUME

Nous proposons ici une représentation des variables à l'issue de l'analyse des correspondances d'un tableau disjonctif complet. En effet, dans les programmes classiques, seules les modalités des variables apparaissent, mais la variable elle-même, prise globalement, n'apparaît pas.

Pour cela, nous associons à chaque variable V un opérateur Ω_V, l'opérateur de projection orthogonale sur le sous espace des indicatrices de la variable V, et nous associons à chaque facteur F l'opérateur de projection orthogonale sur F, Ω_F. L'espace des opérateurs étant muni d'une métrique euclidienne, nous pouvons projeter Ω_V sur les opérateurs associés aux facteurs.

Cette représentation met en évidence, à la fois les liaisons entre une variable et un facteur et les liaisons entre les variables. En effet, la coordonnée de l'opérateur Ω_V sur l'opérateur Ω_F est proportionnelle à la contribution de l'ensemble des modalités de la variable V à l'inertie de F. Et, le cosinus de l'angle entre les opérateurs Ω_V et $\Omega_{V'}$, associés à deux variables V et V' est une mesure de la liaison entre V et V'.

Un exemple commenté illustre la technique que nous proposons.

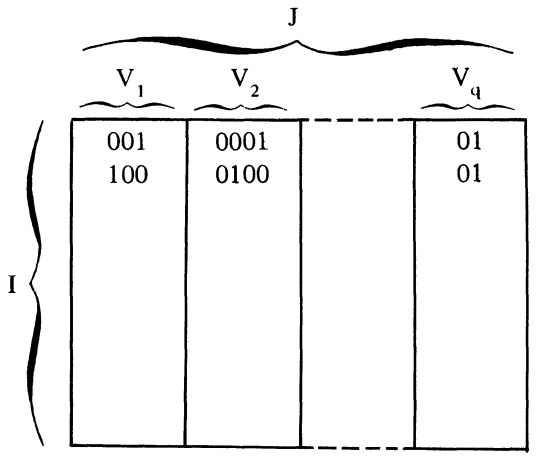
I – LES DONNEES, LE TABLEAU DISJONCTIF COMPLET ET SON ANALYSE (cf. [1] et [4])

Pour analyser un ensemble de variables définissant chacune une partition d'un ensemble d'individus, on lui associe souvent un tableau disjonctif complet que l'on traite par l'analyse des correspondances. Ces variables peuvent être des variables qualitatives à plusieurs modalités, des questions avec plusieurs réponses possibles ou bien encore des variables quantitatives codées par classe.

Le tableau disjonctif complet associé à l'ensemble de variables $V_1, \ldots, V_q$ est le tableau des indicatrices des partitions définies par ces q variables. Il croise donc l'ensemble des individus I, et l'ensemble J des modalités des q variables. Il se compose uniquement de 0 et de 1. La valeur 1 au croisement de la j-ième colonne indique que l'individu i a la modalité j. Dans le groupe de colonnes associé aux modalités d'une même variable, il y a sur une ligne un 1 et un seul. Le tableau a donc la forme suivante :

* Département de Mathématiques – I.N.S.A., B.P. 14A – 35031 RENNES CEDEX.

Mots-clés : **Correspondances multiples – Disjonctif complet – Liaison – Opérateur – Projection.**



Dans l'analyse des correspondances de ce tableau disjonctif complet, on obtient des couples de facteurs définis respectivement sur l'ensemble I des individus et l'ensemble J des modalités de toutes les variables. Mais, les variables elles-mêmes n'apparaissent pas dans les résultats.

Ceux qui appliquent l'analyse des correspondances aux tableaux disjonctifs complets savent que pour interpréter les résultats il est nécessaire de connaître les relations entre un facteur et une variable. Comme ils ne disposent pas de techniques systématiques, ils font quelques calculs simples. Par exemple, ils mesurent l'importance d'une variable V pour un facteur F en calculant la contribution de l'ensemble des modalités de V à l'inertie de F. Ou bien, puisqu'une variable est une partition de l'ensemble I en classes, ils calculent, pour ces classes l'inertie interclasse du facteur F.

Jusqu'à présent les résultats de l'analyse des correspondances n'étaient pas utilisés pour estimer les liaisons entre les variables, seules les liaisons entre les modalités de ces variables étant mises en évidence.

II – OPERATEUR Ω_V ASSOCIE A LA VARIABLE V
(cf. [2] et [5])

Les variables que nous considérons définissent une partition de l'ensemble I. La donnée d'une partition de I en s éléments est équivalente à la donnée des s fonctions indicatrices de cette partition. Ces s fonctions sont des éléments de l'espace vectoriel des fonctions numériques définies sur I, espace noté $\mathbf{R}^I$. Elles engendrent un sous espace vectoriel $\mathcal{E}_V$ de $\mathbf{R}^I$ de dimension s. Et ce sous espace $\mathcal{E}_V$ détermine les s fonctions indicatrices.

La somme de toutes les fonctions indicatrices est la fonction constante sur I, égale à 1. Le sous espace $\mathcal{E}_V$ contient donc cette fonction notée 1_Δ. Il suffit donc de connaître le sous espace E_V de $\mathcal{E}_V$ orthogonal à 1_Δ pour connaître la partition de I. La notion d'orthogonalité dans $\mathbf{R}^I$ est celle de la métrique usuelle.

Un sous espace est caractérisé par l'opérateur de projection orthogonale sur ce sous espace. A une variable V on peut donc associer par une application injective, l'opérateur de projection orthogonale sur le sous espace E_V que nous venons de définir. Nous notons cet opérateur Ω_V.

III – OPERATEUR Ω_F ASSOCIE AU FACTEUR F

Un facteur F sur I est un élément de l'espace $\mathbf{R}^I$, il engendre donc un sous espace $\{F\}$ de $\mathbf{R}^I$ de dimension 1. A un facteur F, nous associons l'opérateur de projection orthogonale sur ce sous espace, et nous noterons cet opérateur Ω_F.

IV – PRODUIT SCALAIRE DE DEUX OPERATEURS SYMETRIQUES

La trace du composé de deux opérateurs symétriques définit un produit scalaire sur l'espace des opérateurs symétriques de $\mathbf{R}^I$ (cf. [2]).

Rappelons que si E et F sont deux sous espaces de $\mathbf{R}^I$, le produit scalaire des opérateurs de projection orthogonale sur ces sous espaces est la somme des cosinus carrés des angles canoniques entre E et F. Le carré de la norme d'un opérateur de projection est donc égal à la dimension de son image.

V – PROJECTION DE $\Omega_V/\|\Omega_V\|$ SUR Ω_F. LIENS ENTRE F ET V

Nous proposons de projeter les opérateurs normés associés aux variables sur les opérateurs Ω_F. Nous montrons dans ce paragraphe que la coordonnée de $\Omega_V/\|\Omega_V\|$ sur Ω_F met en évidence les liens entre F et V.

Puisque Ω_F et Ω_V sont des projecteurs, leur produit scalaire est égal à la somme des carrés des corrélations canoniques entre l'espace $\{F\}$ et l'espace E_V. La dimension de $\{F\}$ étant égale à 1, il y a un seul coefficient de corrélation canonique qui est le cosinus de l'angle θ entre F et E_V.

La dimension de E_V étant égale à s-1, où s est le nombre de modalités de V, on a $\|\Omega_V\| = \sqrt{s-1}$. La projection de l'opérateur normé $\Omega_V/\|\Omega_V\|$ sur Ω_F est donc égale à :

$$\cos(\Omega_F, \Omega_V) = \frac{\langle \Omega_F, \Omega_V \rangle}{\|\Omega_F\| \, \|\Omega_V\|} = \frac{\cos^2\theta}{\sqrt{s-1}}$$

puisque F étant de dimension 1, $\|\Omega_F\| = 1$.

Relation avec l'inertie des modalités de la variable

Dans l'analyse factorielle des correspondances du tableau disjonctif complet, le nuage de points représentant les modalités des variables est situé dans l'espace $\mathbf{R}_I$, dual de $\mathbf{R}^I$. Ces deux espaces sont munis de la métrique usuelle aux coefficients 1/cardI et cardI près. Ces métriques induisent un isomorphisme entre $\mathbf{R}^I$ et $\mathbf{R}_I$.

L'image du facteur F par cet isomorphisme est un axe d'inertie f du nuage des modalités. Et l'image d'une fonction indicatrice de la variable V est un vecteur proportionnel au vecteur joignant l'origine à la modalité en question de la variable V. Le sous espace $\mathcal{E}_V$ des fonctions indicatrices de V a donc pour image le sous espace engendré par l'origine et les modalités de V.

L'angle entre F et $\mathcal{E}_V$ est donc égal à l'angle entre l'axe d'inertie f associé à F et le sous espace engendré par les modalités de V. Or l'angle entre F et $\mathcal{E}_V$ est aussi égal à l'angle θ entre F et E_V puisque F est orthogonal au sous espace 1_Δ des fonctions constantes sur I.

Les vecteurs joignant l'origine aux modalités de V sont orthogonaux. Si q est le nombre total de variables étudiées, l'inertie, par rapport à l'origine, de chaque modalité est 1/q. Donc, l'inertie des projections des modalités de la variable V dans une direction quelconque du sous espace qu'elles engendrent est constante et égale à 1/q. Dans une direction quelconque de $\mathbf{R}_I$, elle sera égale au produit par 1/q du cosinus carré de l'angle entre cette direction et leur sous espace.

Pour la direction f, ce sera $\cos^2\theta/q$. La contribution de la variable V à l'inertie du facteur F est donc, en pourcentage de l'inertie totale λ de F, $\cos^2\theta/(\lambda q)$. Ce pourcentage se calcule facilement avec les résultats de l'analyse des correspondances, car les pourcentages d'inertie fournis par chaque modalité figurent sur les listings. Ceci permet de calculer $\cos^2\theta$ et la coordonnée de $\Omega_V/\|\Omega_V\|$ sur Ω_F.

Cette coordonnée est liée aussi au pourcentage, extrait par F, de l'inertie de l'ensemble des modalités de V. L'inertie de ces s modalités, par rapport au centre de gravité du nuage – qui est aussi leur centre de gravité – est égale à (s – 1), puisqu'elles engendrent un sous espace de dimension s – 1. Le pourcentage d'inertie extrait par la projection sur f est donc $\cos^2\theta/(s-1)$.

Autres interprétations

La variable V définit des classes sur I. Pour ces classes, l'inertie interclasse du facteur, quotienté par son inertie totale est exactement $\cos^2\theta$. Lorsque $\cos^2\theta$ vaut 1, l'inertie interclasse est égale à l'inertie totale, tous les éléments d'une même classe sont confondus sur le facteur F. Donc, plus la coordonnée de $\dfrac{\Omega_V}{\|\Omega_V\|}$ s'approche de $1/\sqrt{s-1}$ plus les éléments de ces classes se retrouvent groupés. Notons que $\cos^2\theta$ est aussi le rapport de corrélation de F par rapport à V.

En résumé

$$\text{Coordonnée de } \frac{\Omega_V}{\|\Omega_V\|} \text{ sur } \Omega_F = \cos(\Omega_V, \Omega_F)$$

$$= \cos^2\theta/\sqrt{s-1}$$

$$= \text{Contribution de V à F} \times \frac{q}{\sqrt{s-1}}$$

$$= \text{Pourcentage d'inertie de V extrait par F} \times \sqrt{s-1}$$

$$= \frac{1}{\sqrt{s-1}} \times \frac{\text{Inertie interclasse de F}}{\text{Inertie de F}}$$

VI – LIAISON ENTRE 2 VARIABLES. SA REPRESENTATION SUR LES GRAPHIQUES

Rappelons d'abord quelques mesures classiques de la liaison entre deux variables définissant chacune une partition d'un même ensemble I. Cette liaison est décrite par le tableau de contingence croisant les deux positions.

Elle est généralement mesurée par le Ψ^2 de ce tableau de contingence, ou le ϕ^2 qui est égale au χ^2 divisé par le cardinal de l'ensemble I. Ou bien encore par le T^2 de Tchuprov qui, si les variables V et V' ont respectivement s et t modalités, vaut $\phi^2/\sqrt{(s-1)(t-1)}$.

Ce dernier, le T^2 est compris entre 0 et 1. Il vaut 1 si, et seulement si, les deux partitions sont identiques. Le tableau de contingence est alors carré et diagonal. Il vaut 0 lorsque ces deux variables sont indépendantes. Plus T^2 est grand, plus les variables sont dépendantes. Le T^2 mesure la dépendance de deux variables qualitatives, comme le coefficient de corrélation linéaire mesure la liaison entre deux variables quantitatives.

Dans l'espace des opérateurs, on retrouve le φ^2 et le T^2 des deux variables V et V'. le produit scalaire des opérateurs Ω_V et Ω_V, associés à V et V' est égal à leur ϕ^2. Et le produit scalaire des opérateurs normés $\frac{\Omega_V}{||\Omega_V||}$ et $\frac{\Omega_{V'}}{||\Omega_{V'}||}$ est égal à leur T^2.

Les opérateurs associés à deux facteurs distincts sont orthogonaux. En effet, la métrique induite sur R^I par un tableau disjonctif complet est, à un coefficient près, la métrique usuelle. Deux facteurs distincts sont donc orthogonaux pour la métrique usuelle et les composés des opérateurs qui leur sont associés sont donc nuls.

La projection sur deux opérateurs Ω_F et Ω_F, associés à deux facteurs F et F' donne donc la projection sur le plan qu'ils engendrent. Nous obtenons ainsi une représentation graphique plane des opérateurs associés aux variables.

Comme il y a une bijection entre les facteurs et les opérateurs qui leur sont associés, on peut représenter directement les variables sur les graphiques de l'analyse des correspondances. Mais, nous proposons plutôt de tracer des graphiques annexes car la projection des modalités d'une variable et la projection de la variable elle-même sont de natures différentes puisqu'elles se font dans des espaces différents.

En résumé, le cosinus de l'angle entre les opérateurs normés représentant les variables V et V' dans l'espace Σ des opérateurs symétriques est égal au T^2 entre ces deux variables. Et nous projetons ces opérateurs sur les opérateurs associés aux facteurs qui forment une base incomplète orthonormée de Σ.

Si deux variables V et V' sont bien représentées par leur projection, l'angle entre les opérateurs Ω_V et $\Omega_{V'}$, et donc le T^2 entre les variables V et V' sera visible directement sur les graphiques, de la même façon qu'en analyse en composantes principales, on voit la corrélation entre deux variables quantitatives bien représentées sur les composantes principales.

La qualité de représentation d'une variable sur un sous espace est mesurée par la longueur de sa projection, puisque l'opérateur associé que l'on projette est normé.

Mais, nous avons montré au paragraphe précédent que la coordonnée de $\frac{\Omega_V}{||\Omega_V||}$ sur Ω_F était égale à $\cos^2\theta/\sqrt{s-1}$. Il faudra donc considérer un nombre de facteurs au moins égal à $s-1$ pour que l'opérateur $\frac{\Omega_V}{||\Omega_V||}$ puisse être bien représenté. Ceci s'explique bien, puisque l'opérateur associé à un facteur est de rang 1 alors que l'opérateur associé à la variable est de rang $s-1$. Mais cela implique que pour des variables ayant plus de 3 modalités, nous devrons consulter plusieurs graphiques plans pour juger de leur proximité.

VII – LIAISON ENTRE UNE VARIABLE ET UN PLAN FACTORIEL

On peut se demander si la représentation de la variable V dans le plan $\{\Omega_F, \Omega_{F'}\}$ permet aussi de visualiser une liaison entre V et le plan factoriel $\{F, F'\}$. C'est ce que nous étudions ici.

La liaison entre V et $\{F, F'\}$ peut se mesurer par des pourcentages d'inertie : le pourcentage, extrait par le plan $\{F, F'\}$ de l'inertie des s modalités de V qui vaut $\frac{\cos^2\theta + \cos^2\theta'}{s-1}$; ou bien, le pourcentage d'inertie fourni par ces modalités à l'inertie du plan $\{F\ F'\}$, qui vaut $\frac{\cos^2\theta + \cos^2\theta'}{q(\lambda+\lambda')}$; ou bien encore, par le pourcentage d'inertie interclasse – pour les classes de V – dans l'inertie du plan $\{F, F'\}$: $\frac{\lambda\cos^2\theta + \lambda'\cos^2\theta'}{\lambda+\lambda'}$.

Toutes ces notions font intervenir plus ou moins directement la somme des coordonnées, $\frac{\cos^2\theta}{\sqrt{s-1}}$ et $\frac{\cos^2\theta'}{\sqrt{s-1}}$, de V. Elles n'apparaissent donc pas sur le graphique plan.

La liaison entre V et $\{F, F'\}$ peut aussi être mesurée par la liaison entre de sous espace E_V et le sous espace $\{F, F'\}$. Donc, par le cosinus de l'angle entre les opérateurs Ω_V et l'opérateur de projection orthogonale sur $\{F, F'\}$. L'opérateur de projection orthogonale sur $\{F, F'\}$ est la somme des opérateurs Ω_F et $\Omega_{F'}$. Il est donc situé dans le plan $\{\Omega_F, \Omega_{F'}\}$, sur la première bissectrice. Le cosinus de l'angle entre Ω_V et cet opérateur est donc égal à la longueur de la projection de Ω_V sur la première bissectrice du plan.

En conclusion, certains liens entre une variable V et le plan $\{F, F'\}$ sont visibles sur les graphiques, mais ce ne sont sans doute pas les plus faciles à interprèter.

VIII – REMARQUES

La description des variables qualitatives est tout à fait analogue à celle qui est obtenue dans [2] et [5] par une analyse en composantes principales des opérateurs associés aux variables.

Mais les opérateurs sur lesquels on projette les opérateurs représentant les variables ne sont pas les mêmes. Dans [2] et [5], ce sont les composantes principales obtenues en diagonalisant la matrice de T^2. Alors qu'ici, ce sont les opérateurs associés aux facteurs de l'analyse des correspondances. Le but de cette technique n'est pas de décrire au mieux les T^2, mais d'apporter une aide à l'interprétation des résultats de l'analyse des tableaux disjonctifs complets.

Cette comparaison amène cependant une remarque. Dans l'analyse en composantes principales des opérateurs, la première composante C maximise la quantité :

$$\sum_{i=1}^{q} \cos^2 (\Omega_{V_i}, C)$$

L'opérateur attaché au premier facteur de l'analyse des correspondances possède aussi une propriété extrémale. En effet, F est le vecteur de $\mathbf{R}^I$ maximisant :

$$\sum_{i=1}^{q} \cos^2 \theta_i$$

où θ_i est l'angle entre F et le sous-espace associé à la variable V_i. Donc, l'opérateur Ω_F associé à F est l'opérateur symétrique de rang 1 maximisant la quantité :

$$\sum_{i=1}^{q} \cos (\Omega_{V_i}, \Omega_F) \times \sqrt{s_i - 1}$$

IX – EXEMPLE

L'exemple que nous avons choisi est un tableau disjonctif complet de 15 variables quantitatives transformées en variables qualitatives par division en 4 classes. Ces variables sont diverses mesures prises sur 120 plantes.

Nous donnons dans le tableau 1 les valeurs des 5 premiers facteurs pour les 60 modalités des 15 variables. Les modalités sont désignées par une lettre caractérisant la variable suivi d'un chiffre indiquant la classe de la variable.

Nous avons calculé les "projections des variables" sur les 4 premiers facteurs, à partir de ce tableau et des valeurs propres. Commentons maintenant les graphiques obtenus.

Plan $[F_1, F_2]$

Les variables ayant chacune 4 modalités, la coordonnée d'une variable V sur un facteur est égale à $\cos^2\theta/\sqrt{3}$. L'angle θ est celui défini au §V et $\cos^2\theta$ est égal au pourcentage d'inertie inter-classe.

Chaque coordonnée est donc comprise entre 0 et $1/\sqrt{3}$. Nous avons fait figurer, sur le plan $[F_1, F_2]$ le carré de côté $1/\sqrt{3}$.

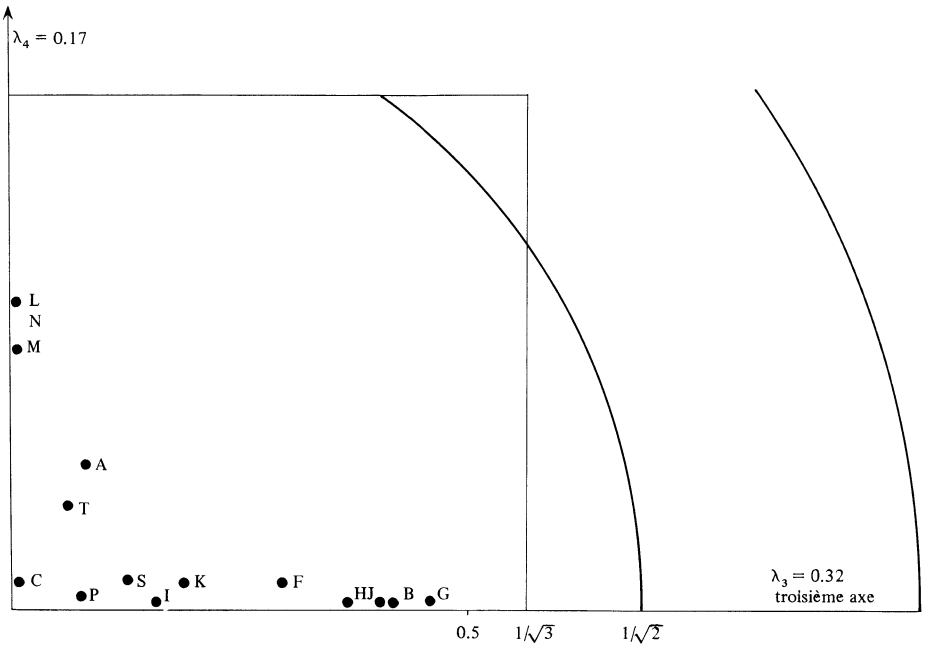
Quand une variable est proche d'un côté extérieur de ce carré, $\cos^2\theta$ est proche de 1. Cette variable est liée au facteur : les éléments de chacune de ses classes sont bien regroupés, le facteur est proche du sous-espace engendré par ses fonctions indicatrices, et le pourcentage d'inertie que leur extrait ce facteur est presque 1/3.

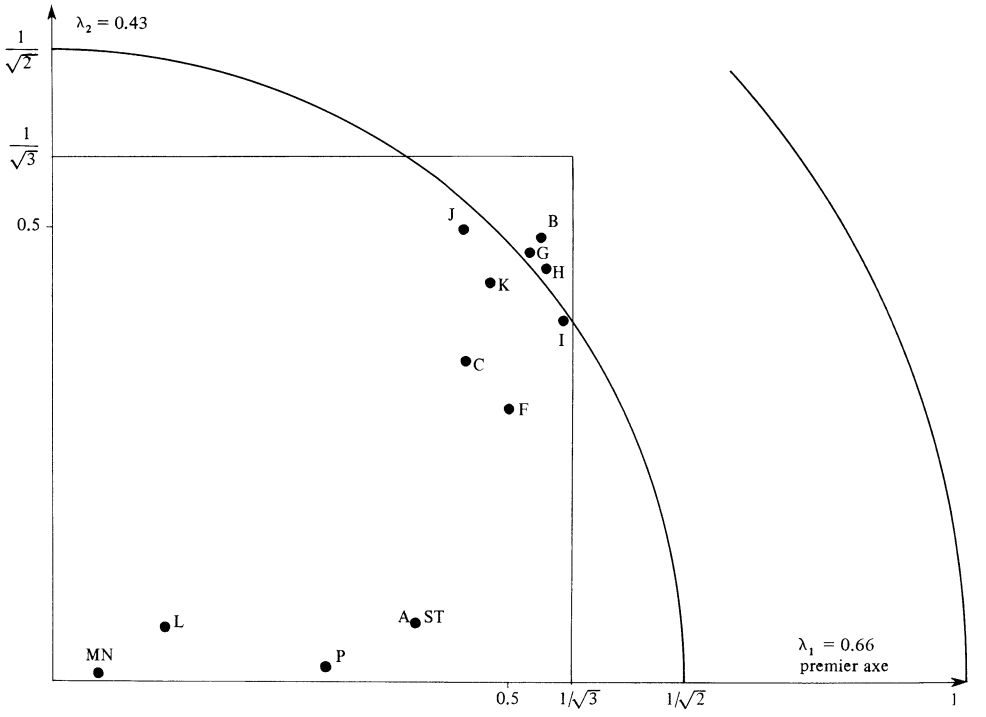
Tableau 1

RESULTATS DE L'ANALYSE DES CORRESPONDANCES :

LES MODALITES DES VARIABLES

J1	QLT	PBID	INRI	1#F	C8R	CTRI	2#F	C8R	CTRI	3#F	C8R	CTRI	4#F	C8R	CTRI	5#F	C8R	CTR
1 A1	730	17	161	1341	637	481	-395	56	61	13	0	01	-243	21	61	212	16	5
2 A2	468	16	171	-15	0	01	501	79	91	288	26	41	252	20	61	-1046	343	119
3 A3	370	15	171	-660	131	101	121	4	11	368	41	71	-771	179	531	-226	15	5
4 A4	805	18	161	-720	192	141	-162	10	11	-582	125	191	675	168	471	917	310	103
5 B1	876	18	161	1385	707	521	-560	116	131	351	45	71	150	8	21	36	0	0
6 B2	800	16	171	156	8	11	843	232	271	-1301	554	871	-76	2	11	-116	4	2
7 B3	801	16	171	-563	100	81	1003	315	371	1064	355	571	111	4	11	295	27	9
8 B4	915	16	171	-1125	414	321	-1200	471	551	-112	4	11	-193	12	41	-207	14	5
9 C1	596	17	161	1223	529	401	-362	47	51	37	0	01	-4	0	01	236	20	7
10 C2	225	16	171	285	25	21	596	111	131	-502	79	131	160	8	21	-71	2	1
11 C3	420	17	171	-447	68	51	878	262	301	487	81	131	-138	7	21	-82	2	1
12 C4	802	16	171	-1112	405	311	-1096	393	461	-54	1	01	-4	0	01	-95	3	1
13 F1	776	17	161	1330	627	471	-522	97	111	273	26	41	274	26	71	37	0	0
14 F2	671	16	171	281	25	21	754	178	211	-1180	437	701	-88	2	11	-301	29	10
15 F3	542	16	171	-578	109	81	703	161	191	775	196	311	-315	33	91	361	43	15
16 F4	649	17	171	-1072	392	301	-851	247	291	77	2	01	108	4	11	-104	4	1
17 G1	857	17	171	1408	674	511	-585	117	131	406	56	91	172	10	31	-19	0	0
18 G2	846	17	161	223	18	11	771	211	241	-1314	612	951	-60	1	01	-101	4	1
19 G3	836	15	171	-584	103	81	1022	314	371	1145	393	631	115	4	11	272	22	8
20 G4	880	17	171	-1105	416	311	-1137	441	511	-91	3	01	-212	15	41	-121	5	2
21 H1	849	17	171	1390	657	501	-570	111	131	393	53	81	270	25	71	-86	3	1
22 H2	756	17	171	251	21	21	781	207	241	-1224	510	801	-112	4	11	-204	14	5
23 H3	795	16	171	-562	99	81	986	305	361	955	286	461	54	1	01	575	104	36
24 H4	897	17	171	-1111	421	321	-1135	439	511	-65	1	01	-207	15	41	-248	21	7
25 I1	806	22	151	1228	719	491	-356	61	61	59	2	01	197	19	51	-97	5	1
26 I2	434	13	181	16	0	01	923	203	251	-940	211	361	-259	16	51	125	4	1
27 I3	512	16	171	-538	91	71	885	245	291	737	170	271	52	1	01	132	5	2
28 I4	806	16	171	-1100	396	301	-1109	402	471	-55	1	01	-104	4	11	-96	3	1
29 J1	828	17	171	1408	674	511	-557	106	121	323	36	61	176	10	31	66	2	1
30 J2	740	17	171	237	19	11	736	185	211	-1232	517	811	-96	3	11	-213	16	5
31 J3	876	16	171	-560	103	81	1041	354	411	1092	389	611	115	4	11	281	26	9
32 J4	936	16	171	-1134	421	321	-1223	489	571	-154	8	11	-195	13	41	-128	5	2
33 K1	776	21	151	1192	655	451	-331	51	51	99	5	11	237	26	71	-290	39	12
34 K2	550	15	171	62	1	01	871	227	271	-771	179	291	-456	63	181	515	80	28
35 K3	559	15	171	-560	94	71	857	220	261	870	227	371	233	16	51	-85	2	1
36 K4	910	15	171	-1168	392	311	-1316	498	601	-241	17	31	-102	3	11	-31	0	0
37 L1	405	17	171	644	141	111	-89	3	01	12	0	01	-796	216	621	365	45	15
38 L2	171	16	171	98	3	01	322	33	41	200	13	21	-555	97	281	-283	25	9
39 L3	202	17	171	-576	113	91	-507	88	101	30	0	01	58	1	01	4	0	0
40 L4	602	17	171	-158	9	11	295	30	31	-229	18	31	1261	541	1541	-102	4	1
41 M1	290	17	161	319	36	31	-61	1	01	0	0	01	-622	137	391	572	116	39
42 M2	89	16	171	262	22	21	108	4	01	-150	7	11	-400	50	151	140	6	2
43 M3	125	13	181	-116	3	01	46	1	01	221	12	21	-325	25	81	-592	84	31
44 M4	580	21	151	-401	72	51	-59	2	01	-19	0	01	1044	484	1281	-224	22	7
45 N1	366	17	161	344	42	31	-90	3	01	121	5	11	-849	256	721	411	60	20
46 N2	95	17	161	167	10	11	-78	2	01	95	3	01	-389	54	151	-267	26	9
47 N3	49	16	171	-347	38	31	72	2	01	-161	8	11	34	0	01	-48	1	0
48 N4	575	16	171	-213	14	11	115	4	01	-74	2	01	1327	551	1611	-107	4	1
49 P1	405	16	171	1062	353	271	-144	7	11	33	0	01	-69	2	01	371	43	15
50 P2	200	18	161	222	19	11	136	7	11	448	77	121	237	22	61	-441	75	25
51 P3	187	13	181	-452	51	41	191	9	11	-592	88	151	-35	0	01	-396	39	14
52 P4	327	19	161	-786	246	181	-144	8	11	-45	1	01	-146	9	21	398	63	21
53 S1	767	16	171	1406	646	491	-472	73	91	114	4	11	-243	20	61	274	24	8
54 S2	405	17	171	-126	5	01	526	94	111	473	76	121	90	3	11	-816	227	77
55 S3	197	16	171	-653	140	111	-108	4	01	180	11	21	-174	10	31	-312	32	11
56 S4	603	17	171	-601	123	91	39	1	01	-758	196	311	317	34	101	855	249	85
57 T1	726	17	161	1304	602	451	-458	75	91	325	37	61	123	5	21	137	7	2
58 T2	440	16	171	-12	0	01	533	93	111	-126	5	11	-240	19	51	-994	323	111
59 T3	264	16	171	-609	121	91	219	16	21	301	30	51	-544	97	281	-37	0	0
60 T4	698	16	171	-761	190	141	-264	23	31	-518	88	141	656	140	401	887	257	88


$\lambda_4 = 0.17$
L
N
M
A
T
C
P
S
I
K
F
HJ
B
G
$\lambda_3 = 0.32$
troisième axe
0.5
$1/\sqrt{3}$
$1/\sqrt{2}$


$\lambda_2 = 0.43$
$\frac{1}{\sqrt{2}}$
$\frac{1}{\sqrt{3}}$
0.5
J
B
G
H
K
I
C
F
L
A
ST
MN
P
$\lambda_1 = 0.66$
premier axe
0.5
$1/\sqrt{3}$
$1/\sqrt{2}$
1

C'est le cas pour le premier facteur pour les variables (I, B, H, G, F, K, C) pour lesquels $\cos^2\theta$ est supérieur à 0,78. C'est aussi le cas, à un degré moindre pour les variables S, T, A, ($\cos^2\theta \cong 0.71$).

Pour le deuxième facteur ceci est encore vérifié par le groupe de variables (J, K, H, B, G).

Par contre, les variables L, M et N sont très proches de l'origine. Ce qui met en évidence leur indépendance avec les 2 premiers facteurs. Il sera donc inutile d'étudier la position de leurs modalités sur les graphiques d'Analyse des correspondances. D'autre part, S, T, A, P sont un peu liées au premier facteur, mais absolument pas au second.

Sur ce graphique, on lit aussi très bien l'importance des variables dans la détermination des facteurs. En effet, les coordonnées des variables sur un facteur sont aussi proportionnelles au pourcentage d'inertie fourni par les variables à ce facteur. Les 2 premiers facteurs sont donc essentiellement déterminés par le groupe de variables situées en haut à droite.

Certaines variables sont très proches sur le plan (1 – 2). Mais cette proximité n'est pas suffisante pour conclure sur leur distance réelle. En effet, le plan des 2 premiers facteurs ne peut suffir pour bien représenter des variables qui ont 4 modalités, et qui sont donc associées à des opérateurs de rang 3. Pour cela, il faudra au moins l'espace des 3 premiers facteurs.

Pour conclure à leur sujet, nous considérerons donc les 2 graphiques des plans $[F_1, F_2]$ et des plans $[F_3, F_4]$. Pour faciliter la lecture, nous avons tracé dans ces deux plans, le cercle de rayon $1/\sqrt{2}$: une variable sera bien représentée dans le sous-espace des 4 premiers facteurs si ses projections dans chacun des deux plans est proche de ce cercle.

Plan $[F_3, F_4]$

Le graphique montre immédiatement que les variables importantes pour le $3^{ème}$ facteur sont les mêmes que celles qui déterminent les 2 premiers.

Et que les variables L, M, N apparaissent au niveau du $4^{ème}$ facteur, et le déterminent quasiment à elles seules. Les variables A et T ne sont pas complètement indépendantes de ce $4^{ème}$ facteur.

D'autre part, on peut maintenant, à l'aide de ces 2 graphiques, affirmer que les variables G, B, J, H qui sont très bien représentées dans $[F_1, F_2, F_3]$ sont réellement très proches. Et que le sous-espace de R^I engendré par leurs 4 modalités est pratiquement le même et très proche du sous-espace $[F_1, F_2, F_3]$.

En conclusion, un simple coup d'oeil sur ces graphiques permet d'orienter le dépouillement de l'analyse. L'interprétation des 3 premiers facteurs devra s'appuyer sur les groupes de variables (G, B, J, H) et aussi (K, I, C, F). Le premier groupe est si homogène qu'il est vraisemblablement inutile d'en différencier les éléments. Il n'est pas étonnant de retrouver sur les graphiques de l'analyse un effet Guttmann dans les plans 1 – 2 et 1 – 3 pour les variables qui sont concernées.

Par contre, le $4^{ème}$ facteur est indépendant des 3 premiers, on doit donc l'interpréter séparément en se basant sur les variables L, M et N.

Vraisemblablement, une étude poussée des graphiques aurait permis d'aboutir aux même conclusions, mais après un long travail.

BIBLIOGRAPHIE

[1] BENZECRI et collaborateurs. – "L'analyse des données" – Dunod, 1973.

[2] CAILLIEZ F. PAGES J.P. – "Introduction à l'analyse des données". SMASH, 1976,

[3] LEBART L. et FENELON J.P. – "Statistique et informatique appliquées". Dunod, 1973.

[4] LEBART L. MORINEAU A., TABARD N. – "Technique de la description statistique".

[5] PAGES J.P., ESCOUFIER Y. CAZES P. – "Opérateur et analyse des tableaux de plus de deux dimensions". *Cahier du B.U.R.O.*, 1976.