DECOMPOSITION CANONIQUE D'UN OPERATEUR ET SON APPLICATION EN ANALYSE DES DONNEES

Jérôme MANUCEAU*

* Professeur , U.E.R. de Mathématiques, Université de Provence ,

3,Place Victor Hugo - 13003 MARSEILLE

TABLE DES MATIERES

4. ANALYSE CANONIQUE (A.C.)

5. PARAMETRES QUALITATIFS ANALYSE FACTORIELLE DISCRIMINANTE (A.F.D.)

6. INDICE DE CORRELATION ANALYSE FACTORIELLE DE CORRESPONDANCES(A.F.C)

Je remercie ,Messieurs Axel GRORUD et Gérard RAUZY pour la lecture critique du manuscrit, qu'ils ont faite .

RESUME .

Nous donnons un critère appelé "critère de non dégénérescence" pour qu'un système statistique admette une représentation . Nous montrons que deux représentations " isomorphes ", d'un même système statistique non dégénéré, sont multiples l'une de l'autre .

La décomposition canonique d'un opérateur, étudiée dans le premier chapitre et appliquée à une représentation, montre que

$$u = \sum_{i=1}^{r} \sqrt{\lambda_i}\ \varepsilon_i \otimes \psi_i$$

où les λ_i sont les valeurs propres de la matrice d'inertie, les$(\varepsilon_i)_i$ sont les" vecteurs principaux" et les $(\psi_i)_i$ sont les "composantes principales".

Nous définissons la notion de " somme de deux représentations " d'un même système statistique et montrons que toute représentation se décompose en une "somme directe" de représentations "irréductibles" .

Nous faisons une Analyse Canonique de deux sous espaces vectoriels d'un espace euclidien et l'appliquons à deux représentations. Nous montrons aussi qu'une représentation comportant des paramètres non centrés, n'est jamais équivalente à la représentation correspondant aux paramètres centrés.

En Analyse Factorielle Discriminante nous décrivons les centres de gravité "intraclasse" et donnons des coefficients représentant la qualité de cette description .

Dans le dernier chapitre nous définissons un "indice de corrélation " entre deux paramètres , compris entre 0 et 1 . Cet indice est nul si et seulement si les paramètres sont indépendants et égal à 1 si et seulement si les paramètres sont égaux .

1. DECOMPOSITION CANONIQUE D'UN OPERATEUR

1.1. Notations et position du problème.

1.1.1. Soit E et F deux espaces euclidiens de dimension ι et n respectivement, u une application linéaire de F dans E, u' sa transposée, c'est-à-dire,

$$\forall x \in E,\ \forall y \in F,\ \langle x \mid u\,y \rangle = \langle u'\,x \mid y \rangle .$$

Posons $V = u\,u'$ et $U = u'u$. V s'appellera "opérateur d'inertie ".

$$\begin{tikzcd}
E \arrow[d, shift left, "I"] & F \arrow[l, "u"'] \arrow[d, shift left, "U"] \\
E \arrow[u, shift left, "V"] \arrow[r, "u'"] & F \arrow[u, shift left, "I"]
\end{tikzcd}$$

Nous allons établir, par la méthode des moindres carrés, la décomposition spectrale de V et de U, dont l'existence est connue. Nous chercherons ensuite une "écriture " de u et u' qui " ressemble " à une décomposition spectrale et que nous appellerons "décomposition canonique ".

On peut admettre l'existence de la décomposition spectrale de V et passer directement à 1.2.8.

1.1.2. Soit $f_1, f_2, \dots, f_n$ une base orthonormale quelconque de F et notons $u_i = u(f_i)$. On appelle " indice d'inertie de la famille $(u_i)_{i \in J_n}$

par rapport à O" , l'expression

$$I_0 = \sum_{i=1}^{n} \|u_i\|^2$$

I_0 représente une "dispersion" des $(u_i)_{i\in J_n}$ par rapport à O

Notation : $\forall k\in\mathbb{N}^*$, $J_k = \{ 1,2,\dots, k \}$.

1.1.3. Si W est un sous-espace vectoriel de E, la projection orthogonale de la famille $(u_i)_{i\in J_n}$ dans W, aura un indice d'inertie noté $I_{o,W}$

$\forall i\in\{1,2,\dots,n\}= J_n$, $u_i= \alpha_i+\beta_i$ où $\alpha_i\in W$ et $\beta_i\in W^{\perp}$

$$I_0 = \sum_i \|u_i\|^2= \sum_i(\|\alpha_i\|^2 + \| \beta_i\|^2) = I_{o,W} + I_{o,W^\perp} .$$

$I_{o,W^\perp}$ représente une "distance" des $(u_i)_{i\in J_n}$ à W. Cette distance est d'autant plus petite que la "dispersion" $I_{o,W}$ est plus grande .

1.1.4. La projection des $(u_i)_{i\in J_n}$ dans W, entraîne une perte d"information " sur les $(u_i)_i$.

Le "degré de fidélité" de cette projection, peut-être repéré par le coefficient $I_{o,W}/I_o$ appelé " part d'inertie expliquée par W ".

1.1.5. En statistique

$$I_0 = \sum_{i=1}^{n} p_i \|x_i\|^2, \text{ où } p_i> 0 \text{ et } \sum_{i=1}^{n} p_i = 1 .$$

Les coefficients p_i ne jouant aucun rôle dans ce chapitre, nous les supprimons. Ils seront remis à partir du chapitre suivant .

1.2. **Décomposition spectrale de V.**

1.2.1. **Proposition**

$$I_0 = Tr(V) = Tr(U).$$

Dém. $I_0 = \sum_i < u_i | u_i > = \sum_i < u(f_i) | u(f_i) >$

$$= \sum < f_i | U f_i > = Tr(U).$$

Remarque : I_0 ne dépend pas de la base choisie dans F. C'est donc une caractéristique de u.

1.2.2. **Proposition**

$\forall x \in E$, si $[x]$ *est le sous-espace vectoriel engendré par* x , *alors,*

$$I_{0,[x]} = \frac{< x | V x >}{< x | x >}$$

Dém. Soit $P_{[x]}$ le projecteur orthogonal sur $[x]$.

$$P_{[x]}(u_i) = < u_i | \frac{x}{\| x \|} > . \frac{x}{\| x \|}$$

$$I_{0,[x]} = \sum_i^n < u_i | \frac{x}{\| x \|} >^2 = \sum_i < u(f_i) | \frac{x}{\| x \|} >^2$$

$$= \sum_i < f_i | \frac{u'(x)}{\| x \|} >^2 = \frac{\| u'(x) \|^2}{\| x \|}$$

$$= \frac{< u'(x) | u'(x) >}{< x | x >} = \frac{< x | V x >}{< x | x >}$$

1.2.3. **Proposition**

$\exists x \in E$ *tel que* $I_{0,[x]}$ *soit maximum.*

Dém. l'application

$$x \in E \quad \to \quad I_{o,[x]} \in \mathbb{R}$$

est continue sur la sphère unité car

$| \, I_{o,[x']} - I_{o,[x'']} \, | \; = | < x' | Vx' > - < x'' | Vx'' > |$

$\leqslant | < x' | Vx' > - < x' | Vx'' > | + | < x' | Vx'' > - < x'' | Vx'' > |$

$\leqslant | < x' | V(x'-x'') > | + | < x'-x'' | Vx'' > |$

$\leqslant 2. \| V \| . \| x'-x'' \|$.

Elle y atteind donc son maximum .

1.2.4. Proposition

$I_{o,[x]}$ *est maximum si et seulement si ,* x *est le vecteur propre de* V *correspondant à sa plus grande valeur propre.*

Dém. d'après 1.2.2. $I_{o,[x]} = \dfrac{< x | Vx >}{< x | x >}$,

donc $I_{o,[Vx]} = \dfrac{< Vx | V^2 x >}{< Vx | Vx >} = \dfrac{< x | V^3 x >}{< x | V^2 x >}$.

Or $< x | Vx >^2 \leqslant < x | x > . < x | V^2 x >$,

d'où $\dfrac{< x | Vx >}{< x | x >} \leqslant \dfrac{< x | V^2 x >}{< x | Vx >}$.

Comme $V \geqslant 0$, le produit scalaire $< x | y >_V = < x | Vy >$ est positif

d'où $< x | Vx >_V{}^2 \leqslant < x | x >_V \; < Vx | Vx >_V$

et $< x | V^2 x >^2 \; \leqslant \; < x | Vx > < x | V^3 x >$;

donc $$\frac{< x \mid V^2 x >}{< x \mid Vx >} \leqslant \frac{< x \mid V^3 x >}{< x \mid V^2 x >}\,.$$

$$I_{0,[x]} = \frac{< x \mid Vx >}{< x \mid x >} \leqslant \frac{< x \mid V^2 x >}{< x \mid Vx >} \leqslant \frac{< x \mid V^3 x >}{< x \mid V^2 x >} = I_{0,[Vx]}$$

D'après 1.2.3., $\exists\, x \in E$ tel que $I_{0,[x]}$ soit maximum. Pour cet x on a donc

$$I_{0,[x]} = I_{0,[Vx]}$$

et $$< x \mid Vx >^2 = < x \mid x > < Vx \mid Vx >$$

ce qui établit la proposition.

1.2.5. Lemme.

Si W_{k+1} *est un sous-espace vectoriel de dimension* $k+1$ *tel que* $I_{0,W_{k+1}}$ *soit maximum,*

alors, $\exists\, W_k \subset W_{k+1}$ *tel que* $\dim(W_k) = k$ et I_{0,W_k} *maximum.*

Dém. Soit W'_k tel que $\dim W'_k = k$ et I_{0,W'_k} maximum.

$$W_{k+1} \cap W_k'^{\perp} \neq \{0\}.$$

Prenons $y \in W_{k+1} \cap W'^{\perp}_k$, $y \neq 0$ et posons $W_{k+1} = [y] \oplus W''$. Alors,

$$I_{0,W_{k+1}} = I_{0,[y]} + I_{0,W''} \geqslant I_{0,W'_k \oplus [y]} = I_{0,W'_k} + I_{0,[y]}$$

donc $I_{0,W''} = I_{0,W'_k}$ puisque $\dim W'' = k$, et W'' est le W_k cherché.

1.2.6. Théorème

$\forall k \leqslant \iota$, $\forall\, W_k$ *sous-espace vectoriel de dimension* k , *alors,*

I_{o,W_k} *est maximum si et seulement si ,* W_k *est somme de sous-espaces propres de* V *correspondant aux valeurs propres les plus grandes .*

Dém. raisonnons par récurrence sur k .

1/ $k = 1$ est vrai (1.2.4.).

2/ Supposons la propriété vraie pour $k < \iota$.

3/ Démontrons-la pour $k+1$. Supposons que W_{k+1} soit de dimension $k+1$ et que $I_{o,W_{k+1}}$ soit maximum.

D'après 1.2.5., $\exists\, W_k \subset W_{k+1}$ tel que dim $W_k = k$ et I_{o,W_k} maximum. En utilisant l'hypothèse de récurrence , il ne reste plus qu'à montrer que $W_k^{\perp} \cap W_{k+1}$ est un sous-espace propre de V .

Prenons $x \in W_k^{\perp} \cap W_{k+1}$, $x \neq 0$ et montrons que $\forall y \in W_k^{\perp}$, $I_{o,[x]} \geqslant I_{o,[y]}$.

En effet $I_{o,W_{k+1}} = I_{o,W_k \oplus [x]} = I_{o,W_k} + I_{o,[x]} \geqslant I_{o,W_k \oplus [y]} = I_{o,W_k} + I_{o,[y]}$.

Donc $I_{o,,[x]}$ est maximum sur $W_k^{\perp}$. Comme $W_k^{\perp}$ est invariant par V (puisque W_k est invariant), on utilise 1.2.4 .

1.2.7 . Proposition

$\forall\, k \leqslant \iota$, $\exists\, W_k$ de dimension k, tel que I_{o,W_k} soit maximum .

Dém. raisonnons par récurrence .

1/ $k = 1$ est vrai (1.2.4)

2/ Supposons la propriété vraie pour $k < \iota$

3/ Démontrons-la pour $k+1$. L'hypothèse de récurrence affirme l'existence de W_k de dimension k tel que I_{o,W_k} soit maximum. D'après 1.2.6. W_k est invariant par V ; il en est de même de $W_k^{\perp}$. 1.2.3. affirme l'existence de $x \in W_k^{\perp}$ tel que $I_{o,[x]}$ soit maximum sur $W_k^{\perp}$. Montrons que $W_{k+1} = W_k + [x]$ est le sous-espace recherché. En effet, si W' est tel que $\dim W' = k+1$ et $I_{o,W_{k+1}}$ maximum, alors $I_{o,W'} \geqslant I_{o,W_{k+1}}$.

Soit $y \in W' \cap W_k^{\perp}$ et W'' tel que $W' = W'' + [y]$.

On a $I_{o,W'} = I_{o,W''} + I_{o,[y]} \geqslant I_{o,W_k} + I_{o,[x]}$. Or par hypothèse $I_{o,W''} \leqslant I_{o,W_k}$ et $I_{o,[y]} \leqslant I_{o,[x]}$.

D'où $I_{o,W'} = I_{o,W_{k+1}}$.

1.2.8 . Théorème .

V admet une décomposition spectrale, c'est-à-dire,

$\exists E_1, E_2, \dots, E_r$ *sous-espaces vectoriels de* E, *deux à deux orthogonaux,*

$\exists \lambda_1, \lambda_2, \dots, \lambda_r$, *nombres strictements positifs, différents deux à deux, tel que*

$$V = \sum_{i=1}^{r} \lambda_i P_{E_i} \;, \quad E = \Big(\bigoplus_{i=1}^{r} E_1\Big) \oplus \operatorname{Ker} V \,,$$

où P_{E_i} est le projecteur orthogonal sur E_i.

Dém : supposons par commodité $\lambda_1 > \lambda_2 > \dots > \lambda_r > 0$. Ce théorème se déduit aisément de 1.2.6 et 1.2.7.

Si $\dim E_i = \iota_i$, alors $\operatorname{Tr}(V) = \sum_{i=1}^{r} \lambda_i \iota_i$.

1.2.9. **Remarque, notation**.

Posons $\overset{-1}{V} = \sum_{i=1}^{r} \lambda_i^{-1} P_{E_i}$. Alors,

$$\overset{-1}{V} = V^{-1} \text{ si } V \text{ est injectif, et } V\overset{-1}{V} = \overset{-1}{V}V = P_{val(V)} = P_{(\ker V)^{\perp}}$$

1.3 . DECOMPOSITION CANONIQUE DE u .

1.3.1. **Proposition**

ker V = ker u'.

Dém. $\forall\ x \in \ker V$, $\|u'(x)\|^2 = \langle Vx \mid x \rangle = 0$.

Donc $\ker V \subset \ker u'$. L'inclusion inverse est évidente.

1.3.2. **Théorème**.

Avec les notations de 1.2.8., il existe $F_1, F_2, \ldots, F_r$ *sous espaces vectoriels de F, deux à deux orthogonaux ; il existe* $P_1, P_2, \ldots, P_r$, *applications de* F *dans* E *telles que*

$$* \ F = \left(\bigoplus_{i=1}^{r} F_i\right) \oplus \ker u,$$

$$* \ P_{i|F_i} = \frac{1}{\sqrt{\lambda_i}}\, u|_{F_i}$$ *soit un isomorphisme de* F_i *sur* E_i *(conservant le produit scalaire),*

$$* \ P_{i|F_i^{\perp}} = 0$$

$$* \ u \circ P_{F_i} = P_{E_i} \circ u$$

$$* \ u = \sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i$$

Cette décomposition appelée " décomposition canonique de u " est unique.

Dém. posons, F_i sous espace propre de U à la valeur propre λ_i. On voit aisément que

$$u'(E_i) \subset F_i \text{ et } u(F_i) \subset E_i .$$

Comme u et u' sont des isomorphismes entre $(\ker u)^{\perp}$ et $(\ker u')^{\perp}$,

$u'(E_i) = F_i$ et $u(F_i) = E_i$.

On déduit le théorème des égalités suivantes :

$$\forall x \in E_i \ , \ \forall x' \in E_i ,$$

$$< u'(x) | u'(x') > = < x \mid V x' > = \lambda_i < x \mid x' > \ ;$$

$$\forall y \in F_i \ , \ \forall y' \in F_i$$

$$< u(y) \mid u(y') > = < y \mid U y' > = \lambda_i < y \mid y' > .$$

1.3.3. Corollaire

val V = val u, val u' = (ker u)$^{\perp}$

$$u' = \sum_{i=1}^{r} \sqrt{\lambda_i}\, P'_i \ , \quad P'|_{E_i} = (P_{i|F_i})^{-1} \ ,$$

$$U = \sum_{i=1}^{r} \lambda_i P_{F_i} \ .$$

1.3.4. Remarque, notation .

Posons $\overset{-1}{u} = \sum_{i=1}^{r} (\sqrt{\lambda_i})^{-1} P'_i$.

Alors $\overset{-1}{u} = u^{-1}$ si u est injectif , et $\overset{-1}{u} \circ u = P_{\text{val } u'}$, $u \circ \overset{-1}{u} = P_{\text{val } u}$.

2. ESPACES PRINCIPAUX

2.1. Généralités

2.1.1. Indice d'inertie

Supposons comme dans le premier chapitre avoir deux espaces vectoriels E et F de dimension ι et n respectivement et u, une application linéaire de F dans E .

Soit $f_1, f_2, \ldots, f_n$ une base orthonormale de F et $p=(p_1, p_2, \ldots, p_n)$ un n^{uple} de nombres strictement positifs vérifiant

$$\sum_{i=1}^{n} p_i = 1 .$$

On considère l'opérateur

$$D_{\sqrt{p}} = \sum_{i=1}^{n} \sqrt{p_i}\ P_{[f_i]}$$

qui est un automorphisme positif de F vérifiant

$$\forall i \in \{1,2,\ldots,n\} \quad D_{\sqrt{p}}(f_i) = \sqrt{p_i}\ f_i .$$

On utilisera parfois l'opérateur

$$D_p = D^2_{\sqrt{p}} = \sum_{i=1}^{n} p_i\ P_{[f_i]}$$

Quel que soit u linéaire : $F \rightarrow E$ on note

$$u_p = u \circ D_{\sqrt{p}} .$$

L'indice d'inertie des $(u_{p(f_i)})_i$ par rapport à O est :

$$I_O = \sum_{i=1}^{n} \| u_{p(f_i)} \|^2 = \sum_{i=1}^{n} p_i \| u_i \|^2 .$$

C'est la définition prise en statistique pour les $(u_i)_i$; p_i est appelé le " poids " de u_i .

On voit donc que changer la définition de l'indice d'inertie (par rapport au premier chapitre) , revient à remplacer u par u_p ; on peut ainsi utiliser les résultats du premier chapitre pour u_p .

Remarque. Dans la plupart des cas , $\forall i \in J_n$, $p_i = \frac{1}{n}$.

2.1.2. Centre de gravité des (u_i)

On appelle centre de gravité des $(u_i)_i$, le vecteur

$$g = \sum_{i=1}^{n} p_i u_i$$

Si nous posons $\varphi_p = \sum_{i=1}^{n} \sqrt{p_i}\ f_i$, alors ,

$$g = u_p(\varphi_p)$$

2.1.3. Si W est un sous espace vectoriel de E , on appellera " distance des $(u_i)_i$ à W" le nombre $I_{o,W^\perp}$ dans lequel on remplace u par u_p. c'est-à-dire

$$I_{o,W^\perp} = \sum_{i=1}^{n} p_i \| P_{W^\perp}(u_i) \|^2$$

En particulier $I_{o,W^\perp} = 0 \Leftrightarrow \forall\ i \in J_n$, $u_i \in W$.

2.2. Espaces principaux .

2.2.1. Définition .

Pour tout entier k $< \ell$, *un sous espace vectoriel de dimension* k,
de E, sera dit " principal " pour les $(ui)_i$, *si la "distance " (2.1.3) des* $(ui)_i$ *à* W *est minimum.*

2.2.2. Proposition

W est un espace principal des $(u_i)_{i's}$ *, si et seulement si , W est somme de sous espaces propres de* $V_p = u \circ D_p \circ u' = u_p \circ u'_p$ *correspondant aux valeurs propres les plus grandes .*

Dém : puisque $I_o = I_{o,W} + I_{o,W^\perp}$,

$I_{o,W^\perp}$ est minimum lorsque $I_{o,W}$ est maximum et on utilise 1..2.6 .

2.2.3. Intéret des espaces principaux .

En Analyse des Données, on cherche à décrire un ensemble de vecteurs , dans un espace vectoriel. Ce problème étant généralement trop complexe, on projette les vecteurs sur un sous-espace vectoriel W de dimension plus faible (1.1.4). Le " degré de fidélité " de cette projection est repéré comme nous l'avons déjà vu par $I_{o,W}/I_o$ qui est la part d'inertie expliquée par W.

Ce degré de fidélité sera maximum pour une dimension k fixée de W, quand ce dernier sera un espace principal.

La plupart du temps on utilisera en fait les plans principaux, c'est-à-dire, les plans engendrés par les vecteurs propres de V_p correspondants aux valeurs propres les plus grandes.

2.2.4. Si $e_1, e_2, \dots, e_L$ est une base orthonormale de E, les espaces princpaux des $(u'_p(e_j))_j$ sont donnés par les sous-espaces propres de

$$U_p = u'_p \circ u_p = D_{\sqrt{p}} \circ u' \circ u \circ D_{\sqrt{p}} \; ..$$

2.2.5. Remarque.

En Analyse des Données on cherche les sous-espaces principaux.

1/ Des $(u_i)_{i \in J_n}$ dans E, en utilisant pour définition de $I_{o,W}$, 2.1.1. et 2.1.3. On a déjà vu que cela revient à rechercher les espaces principaux des $(\sqrt{p_i}\, u_i)_i$ en utilisant pour définition de $I_{o,W}$, 1.1.2 et 1.1.3 $(\sqrt{p_i}\; u_i = u_p(f_i))$.

2/ Des $(u'_j)_{j \in J_L}$ où $u'_j = u'(e_j)$, dans F, en prenant pour produit scalaire dans F

$$< x \,|\, D_p\, y > .$$

Cela revient à chercher les espaces principaux des $(u'_p(e_j))_j$ en ne changeant pas le produit scalaire de F puisque

$$< u'_p(e_j) \,|\, u'_p(e_k) > \; = \; < u'(e_j) \,|\, D\, u'_p(e_k) > .$$

L'avantage de ce choix est que l'interprétation des distances entre les projections des $(u'_p(e_j))_j$ sur un plan principal , sera plus facile , car il s'agira de la distance enclidienne plus proche de notre intuition .

Le deuxième avantage de ce choix est dans la conclusion qui suit.

2.2.6. Conclusion .

Les espaces principaux des $(u_p(f_i))_{1 \leqslant i < n}$ et des $(u'_p(e_j))_{1 \leqslant j \leqslant l}$ sont donnés par la décomposition canonique de u_p. Ils ne dépendent en aucun cas des bases orthonormales choisies dans F et dans E.

En outre si W_k(resp.W'_k) est l'espace principal de dimension k dans E (resp.F), alors ,

$$u_p(W'_k) = W_k$$

$$u'_p(W_k) = W'_k ,$$

et la part d'inertie expliquée par W_k est la même que celle de W'_k (1.3.2 et 1.3.3).

2.3. Espaces affines principaux .

2.3..1. Proposition

Si nous notons $\varphi = \sum_{i=1}^{n} f_i$ et $\varphi'_p = \Sigma\, p_i f_i$, *alors,*

l'opérateur linéaire $C_{\varphi'_p}$ *définit par*

$\forall\, i \in [\, 1,2,...., n\,]$ $C_{\varphi'_p}(f_i) = f_i - \varphi'_p$

est le projecteur sur $[\,\varphi\,]^{\perp}$ et parallèlement à φ'_p ,

c'est-à-dire , ker $C_{\varphi'_p} = [\,\varphi'_p]$ et val $C_{\varphi_p} = [\,\varphi\,]^{\perp}$.

Dém. il est évident que

$$C_{\varphi'_p}(\varphi'_p) = 0 \text{ , que } C^2_{\varphi'_p} = C_{\varphi'_p}$$

et que $\forall\ \Psi = \sum_{i=1}^{n} \alpha_i f_i$, alors , comme

$$C_{\varphi'_p}(\Psi) = \Psi - (\sum_{i=1}^{n} \alpha_i)\,\varphi'_p = \Psi - <\Psi|\varphi>\ \varphi'_p$$

on a

$$C_{\varphi'_p}(\Psi) = \Psi \iff <\Psi\,|\,\varphi> = 0 .$$

Remarquons que l'opérateur $\varphi'_p \otimes \varphi$ définit par

$$\forall\ \Psi \in F\ ,\ (\ \varphi'_p \otimes \varphi\ )(\ \Psi) = <\varphi\,|\,\Psi>\varphi'_p$$

n'est autre que le projecteur sur φ'_p parallèlement à $[\ \varphi]^{\perp}$ et que

$C_{\varphi'_p} = I - \varphi'_p \otimes \varphi$ est le projecteur complémentaire .

2.3.2. Remarque

Si au lieu de u nous prenons l'opérateur $u \circ C_{\varphi'_p}$, on voit que

$$y_i = (\ u \circ C_{\varphi'_p})(f_i) = u_i - g$$

où $g = u(\varphi'_p) = \sum_{i=1}^{n} p_i\, u_i$ est le centre de gravité des u_i .

L'indice d'inertie des y_i est

$$I_g = \sum_{i=1}^{n} p_i\, \| y_i \|^2 = \sum_{i=1}^{n} p_i\, \| u_i - g \|^2 .$$

2.3.3. Proposition .

La fonction $a \in E \ \rightarrow\ I_a = \sum_{i=1}^{n} p_i\, \|u_i - a\|^2 \in \mathbb{R}$

atteind son minimum au point g.

I_a est appelé , l'indice d'inertie des $(u_i)_i$ par rapport à a .

Dém. en effet

$$I_a = \sum_{i=1}^{n} p_i < u_i - g + g - a\ |\ u_i - g + g - a > = I_g + \| g - a \|^2$$

En particulier, pour a = 0

$$I_g = I_0 - \| g \|^2 \ .$$

2.3.4. Remarque importante

Remplacer u par $u \circ C_{\varphi'_p}$ permet de passer de la famille $(u_i)_i$ à la famille $(y_i)_i = (u_i - g)_i$ dont l'indice d'inertie par rapport à 0 est égal à l'indice d'inertie des $(u_i)_i$ par rapport à g et dont le centre de gravité $\sum_{i=1}^{n} p_i y_i$ est en 0 .

Cette transformation est constamment utilisée en Analyse des Données . Or nous verrons plus loin (4.3.6.) que les systèmes statistiques définis par u et par $u \circ C_{\varphi'_p}$ ne sont pas équivalents .

Nous n'utiliserons donc pas dans la suite cette transformation, puisqu'elle change de façon non équivalente, le système étudié. En outre elle ne simplifie pas les calculs .

Par contre cette transformation permet de résoudre le problème suivant : si $(u_i)_i$ est une famille de vecteurs de E et k un entier $< \iota$, quel est l'espace affine de dimension k qui soit " le plus proche " (au sens des moindres carrés) des $(u_i)_i$. C'est l'objet de la suite de ce paragraphe .

2.3.5. Si W est un sous-espace vectoriel de E, on note $I_{0,W}$ l'indice d'inertie de la projection orthogonale des $(u_i)_i$ sur W , par rapport à la projection orthogonale de a sur W,

$$I_{a,W} = \sum_{i=1}^{n} p_i \| P_W(u_i) - P_W(a) \|^2 .$$

Proposition .

$$I_g = I_{g,W} + I_{g,W^\perp}$$

et en outre , $\forall$ a $\in W^\perp$,

$$I_{g,W+a} = I_{g,W} \cdot$$

Dém. $\forall\, i \in J_n$ $\quad u_i = \alpha_i + \beta_i$ où $\alpha_i \in W$ et $\beta_i \in W^\perp$

$$g = g_W + g_{W^\perp} \text{ où } g_W = P_W(g) \text{ et } g_{W^\perp} = P_{W^\perp}(g).$$

$$I_{g,W} = \sum_{i=1}^{n} p_i \|\alpha_i\|^2 - \|g_W\|^2 = I_{o,W} - \|g_W\|^2$$

La première égalité est évidente puisque

$$\|u_i\|^2 = \|\alpha_i\|^2 + \|\beta_i\|^2$$

et

$$\|g\|^2 = \|g_W\|^2 + \|g_{W^\perp}\|^2.$$

De plus $I_{g,W+a} = \sum_{i=1}^{n} p_i \|\alpha_i + a\|^2 - \|g_W + a\|^2 = I_{g;W}$.

Ainsi la dispersion des $(u_i)_i$ projetés dans W et $W+a$ est la même par rapport à leur centre de gravité respectif.

2.3.6. Nous avons vu dans 2.1.3. que la distance des $(u_i)_i$ à W est $I_{o,W^\perp}$. Nous allons définir la distance des $(u_i)_i$ à $W+a$:

$$u_i = \alpha_i + P_{W^\perp}(a) + \beta_i - P_{W^\perp}(a),$$

on voit que $\alpha_i + P_{W^\perp}(a) \in W+a$ et que

$$u_i - (\alpha_i + P_{W^\perp}(a)) = \beta_i - P_{W^\perp}(a) \in W^\perp.$$

On appellera donc, distance des $(u_i)_i$ à $W+a$ le nombre

$$\sum_{i=1}^{n} p_i \|\beta_i - P_{W^\perp}(a)\|^2 = I_{a,W^\perp}$$

2.3.7. Proposition .

La distance des $(ui)_i$ à $W+a$, quand on fait varier a est minimum pour $a = g$.

Autrement dit parmi tous les espaces affines parallèles à W, c'est celui qui passe par g qui est le plus près des $(u_i)_i$.

Dém. $$I_{a,W^\perp} = I_{o,W^\perp} + \|P_{W^\perp}(a)\|^2 - 2 < g_{W^\perp} | P_{W^\perp}(a) >.$$

Or

$\|P_{W^\perp}(a)\|^2 - 2 < g_{W^\perp} | P_{W^\perp}(a) > = < P_{W^\perp}(a) - 2g_{W^\perp} | P_{W^\perp}(a) > .$

$= < (P_{W^\perp}(a) - g_{W^\perp}) - g_{W^\perp} | (P_{W^\perp}(a) - g_{W^\perp}) + g_{W^\perp} >$

$= \|P_{W^\perp}(a) - g_{W^\perp}\|^2 - \|g_{W^\perp}\|^2$.

D'où

$$I_{a,W^\perp} = I_{0,W^\perp} - \|g_{W^\perp}\|^2 + \|P_{W^\perp}(a) - g_{W^\perp}\|^2$$

qui démontre la proposition .

Remarque . $I_{g,W^\perp}$ est aussi bien la distance des $(u_i)_i$ à W+g que celle des $(y_i)_i = (u_i - g)_i$ à W .

2.3.8. Proposition .

Le sous espace affine de dimension $k < \iota$, *le plus proche des* (u_i) *est W+g , où W est l'espace principal de dimension k donné par la décomposition canonique de l'opérateur*

$$u \circ C_{\varphi'_p} \circ D_{\sqrt{p}} = (u \circ C_{\varphi'_p})_p .$$

Dém. Cette proposition se déduit simplement de 2.3.7.(proposition et remarque) et de 2.2.3.

3. SYSTEME STATISTIQUE (S.S) ET SES REPRESENTATIONS. ANALYSE EN COMPOSANTES PRINCIPALES (A.C.P.)

3.1. Définitions .

3.1.1. Système statistique (S.S..)

Un système statistique est un ensemble fini de fonctions, appelées " paramètres " ou " caractères " ou " variables aléatoires ", appliquant un ensemble fini, appelé

" ensemble des individus ", dans $\mathbb{R}$.

Il s'agit ici de " paramètres quantitatifs ". Le cas des " paramètres qualitatifs " sera étudié ultérieurement.

Nous prendrons comme ensemble des individus $J_n = \{1,2,...,n\}$. Le couple (n, ℓ) où ℓ est le nombre des paramètres et n le nombre des individus sera appelé, " dimension du S.S. ".

Deux S.S. de même dimension, dont les paramètres de l'un sont multiples des paramètres de l'autre, seront considérés comme identiques. I_n est un codage des " individus ". Le S.S. ne change pas si nous permuttons les éléments de J_n.

3.1.2. Représentations.

Une représentation d'un système statistique de dimension (n, ℓ) est un triplet

$$(u, (f_i)_{i\in J_n}, (e_j)_{j\in J_\ell})$$

où

** u est une application linéaire d'un espace vectoriel F de dimension n, dans un espace vectoriel E de dimension ℓ;*

* $(f_i)_{i\in J_n}$ *une base orthonormale de F;*

*$(e_j)_{j\in I_\ell}$ *une base orthonormale de E;*

tel que :

$$u_i^j = \langle e_j \mid u(f_i)\rangle$$

soit la valeur que prend le j^e paramètre sur le i^e individu.

E est appelé, " espace des individus " car $u_i = u(f_i)$ représente le i^e individu ; en effet, ses composantes sur la base $(e_j)_{j\in I_\ell}$ sont les valeurs que prennent les paramètres sur le i^e individus.

F est appelé, " espace des paramètres " car $u_j' = u'(e_j)$ représente le j^e paramètre ; en effet, ses composantes sur la base $(f_i)_{i\in J_n}$ sont les valeurs que prend le j^e paramètre.

Evidemment $u'^i_j = u^j_i$.

3.1.3. Caractéristiques classiques des paramètres.

3.1.3.1. Comme dans le deuxième chapitre, supposons qu'à chaque individu est associé un poids $p_i > 0$, où $\sum_{i \in J_n} p_i = 1$.

Notons $\varphi_p = \sum_{i=1}^{n} \sqrt{p_i}\, f_i$ et $\varphi'_p = \sum_{i=1}^{n} p_i f_i$. On voit que $\|\varphi_p\| = 1$ et que

$$F = [\varphi_p] \oplus [\varphi_p]^{\perp}$$

$$I = P_{[\varphi_p]} + P_{[\varphi_p]^{\perp}} .$$

3.1.3.2. Moyenne, variance, écart-type.

A tout paramètre u'_j , est associé deux indices :

* un " indice de position " appelé " moyenne "

$$m_j = \sum_{i=1}^{n} p_i u'^i_j ,$$

* un " indice de dispersion " appelé " variance "

$$\mathrm{var}(u'_j) = \sum_{i=1}^{n} p_i (u'^i_j - m_j)^2 .$$

On appelle " écart-type " le coefficient

$$\sigma_j = \sqrt{\mathrm{var}\,(u'_j)} .$$

Or si nous prenons comme paramètre $u'_p(e_j)$ au lieu de $u'(e_j)$ (2.1.1.), on voit que

et que $u'(e_j) = P_{[\varphi_p]}(u'_p(e_j)) + P_{[\varphi_p]^{\perp}}(u'_p(e_j))$

$$\begin{aligned} P_{[\varphi_p]}(u'_p(e_j)) &= < u'_p(e_j) \mid \varphi_p > \varphi_p \\ &= < u'(e_j) \mid \varphi'_p > \varphi_p \\ &= (\sum_{i=1}^{n} p_i u'^i_j)\, \varphi_p \\ &= m_j \varphi_p . \end{aligned}$$

D'où $\qquad m_j = < u'_p(e_j) \mid \varphi_p >$

et $\qquad |m_j| = \| P_{[\varphi_p]}(u'_p(e_j)) \|$

En outre

$$P_{[\varphi_p]^\perp}(u'_p(e_j)) = u'_p(e_j) - m_j \varphi_p \, ,$$

$$\| u'_p(e_j) \|^2 = \sum_{i,j} < D_{\sqrt{p}}\, u'^i_j \; f_i \mid D_{\sqrt{p}}\, u'^k_j \; f_k >$$

$$= \sum_i p_i (u'^i_j)^2 \, ,$$

d'où

$$\| P_{[\varphi_p]^\perp}(u'_p(e_j)) \|^2 = \| u'_p(e_j) \|^2 - m_j^2$$

$$= \sum_i p_i (u'^i_j)^2 - m_j^2 = \sum_i p_i (u'^i_j - m_j)^2 .$$

Et on déduit que

$$\mathrm{val}\,(u'_j) = \| P_{[\varphi_p]^\perp}(u'_p(e_j)) \|^2$$

$$\sigma_j = \| P_{[\varphi_p]^\perp}(u'_p(e_j)) \|$$

La décomposition de F en somme directe $[\varphi_p] \oplus [\varphi_p]^\perp$, permet donc de décomposer tout paramètre $u'_p(e_j)$ en une somme de deux vecteurs , dont l'un contient l'information sur la moyenne , et l'autre sur l'écart-type .

3.1.3.3. **Remarque** .

$$P_{[\varphi_p]}(f_i) = \sqrt{p_i}\; \varphi_p$$

$$P_{[\varphi_p]^\perp}(f_i) = (I - P_{[\varphi_p]})(f_i) = f_i - \sqrt{p_i}\; \varphi_p \, .$$

$$(u_p \circ P_{[\varphi_p]^\perp})(f_i) = \sqrt{p_i}\,(u(f_i) - g) .$$

D'où $\quad u_p \circ P_{[\varphi_p]^\perp} = (u \circ C\varphi')_{p\,p} \qquad (2.3.8., \quad 2.3.1)$

3.1.4. Vecteur et fonction de répartition .

Poson $J_{n,j}(s) = \{k \in J_n \mid u'^{\,k}_j \leqslant s\}$. On appellera " vecteur de répartition " de u'_j la fonction de $\mathbb{R}$ dans F définie par

$$R_j(s) = \sum_{k \in J_{n,j}(s)} \sqrt{p_k}\, f_k .$$

Sa moyenne n'est autre que la " fonction de répartition " de u'_j car

$$P_{[\varphi_p]}(R_j(s)) = F_j(s) \quad \varphi_p$$

et

$$F_j(s) = \sum_{k \in J_{n,j}(s)} p_k = \| R_j(s) \|^2 .$$

En outre

$$P_{[\varphi_p]^\perp}(R_j(s)) = R_j(s) - F_j(s)\, \varphi_p$$

et

$$\| P_{[\varphi_p]^\perp}(R_j(s)) \|^2 = F_j(s)\,(1 - F_j(s)) .$$

3.1.5. Covariance

On sait que la covariance de u'_j et $u'_{j'}$ est

$$\mathrm{cov}(u'_j , u'_{j'}) = \sum_{k \in J_n} p_k\, (u'^{\,k}_j - m_j)\ (u'^{\,k}_{j'} - m_{j'})$$

$$= < P_{[\varphi_p]^\perp}(u'_p(e_j) \mid P_{[\varphi_p]^\perp}(u'_p(e_{j'})) >$$

$$= \ < u'_p(e_j) \mid u'_p(e_{j'}) > - m_j\, m_{j'}$$

D'où

$$< u'_p(e_j) | u'_p(e_{j'}) > = \mathrm{cov}\,(u'_p, u'_{j...}) + m_j\, m_{j'} .$$

3.1.6. Opérateur d'inertie

L'opérateur d'inertie :

$$V_p = u_p \circ u'_p = u \circ D_p \circ u'$$

est la matrice de variance-covariance, que si les paramètres sont centrés, c'est-à-dire que leur moyenne est nulle. En effet :

$$(V_p)^j_{j'} = < e_j | V_p e_{j'} > = < u'_p(e_j) | u'_p(e_{j'}) >$$

$$= \mathrm{cov}(u'_j, u'_{j'}) + m_j m_{j'}$$

$$(V_p)^j_j = \| u'_p(e_j) \|^2 = \sigma_j^2 + m_j^2) .$$

Enfin

$$\mathrm{Tr}(V_p) = I_0 = \sum_{j \in I\ell} (\sigma_j^2 + m_j^2) .$$

3.2 Représentations d'un même S.S.

3.2.1. Définition

Deux ensembles de vecteurs, d'un espace euclidien, ayant même nombre d'éléments, $(x_i)_{i \in J_n}$ *et* $(y_i)_{i \in J_n}$ *, seront dit "isomorphes", si et seulement si, il existe une permutation* σ *de* J_n *et un opérateur* $T = \lambda O$*, où* λ *est un scalaire et* O *une isométrie tels que :*

$$y_{\sigma(i)} = T(x_i) \quad , \forall\ i \in J_n .$$

On voit que σ permet de réordonner les $(y_i)_{i \in Jn}$ et que T vérifie:

* $T\,T' = T'\,T = \lambda^2 I$,

* T conserve les angles car ,

$$< \frac{x_i}{\|x_i\|} \mid \frac{x_j}{\| x_j\|} > = < \frac{T x_i}{\|T x_i\|} \mid \frac{T x_j}{\|T x_j\|} > = < \frac{y_{\sigma(i)}}{\|y_{\sigma(i)}\|} \mid \frac{y_{\sigma(j)}}{\|y_{\sigma(j)}\|} >$$

* T conserve les rapports des longueurs car ,

$$\frac{\|x_i\|}{\|x_j\|} = \frac{\|Tx_i\|}{\|Tx_j\|} = \frac{\|y_{\sigma(i)}\|}{\|y_{\sigma(j)}\|}$$

Dorénavant , nous supposerons que $\sigma = I$, et si ce n'est pas le cas , nous remplacerons les $(y_i)_{i\in Jn}$ par les $(z_i)_{i\in Jn}$ où

$$z_i = y_{\sigma(i)}, \quad \forall\, i \in Jn \quad .$$

3.2.2. Définition .

Deux représentations de même dimension

$(u,(fi)_{i\in Jn}, (ej)_{j\in J\ell})$ et $(v,(gi)_{i\in J_n}, (hj)_{j\in J\ell})$

seront dit " isomorphes ", si et seulement si ,

$(u(f_i))_{i\in Jn}$ *est isomorphe à* $(v(g_i)_{i\in Jn}$ *et* $(u'(e_j))_{j\in J\ell}$ *est isomorphes à* $(v'(h_j))_{j\in J\ell}$.

Remarque . Il est clair que deux représentations de même dimension isomorphes , peuvent être considérées, sur le plan algébrique, comme identiques . Par conséquent, l'association que nous avons réalisée, entre S.S. et représentation , ne peut avoir un intéret, que si elle est bijective à un isomorphisme près .

C'est le problème que nous allons examiner dans la suite .

3.2.3. Proposition

Si $(u,(f_i)_{i\in Jn},(e_j)_{j\in J\ell})$ *est une représentation d'un S.S.., alors,*

quelque soit les bases orthonormales $(f'_i)_{i\in Jn}$ *de F et* $(e'_j)_{j\in J\ell}$ *de E,alors, il existe une représentation*

$$(v,(f'_i)_{i\in Jn}, (e'_j)_{j\in J\ell})$$

du même S.S. et isomorphe à la première représentation.

Dém. considérons les isométries définies par $Oe_j = e'_j$ et $N f'_i = f_i$.

Alors, si $v = O \circ u \circ N$,

$$v(f'_i) = (O \circ u \circ N) (f'_i) = O(u(f_i))$$

et $\quad v'(e'_j) = (N' \circ u' \circ O')(e'_j) = N'(u'(e_j))$.

En outre

$$< v'(e'_j) \mid f'_i > = < u'(e_j) | f_i > .$$

Remarque . Cette proposition nous permet de prendre les mêmes bases pour toutes les représentations d'un même S.S. Nous prendrons donc dans la suite $F = \mathbb{R}^n$, $E = \mathbb{R}^{\ell}$, avec leur base canonique respective .

Une représentation d'un S.S. de dimension (n, ℓ) sera une application linéaire $u : \mathbb{R}^n \to \mathbb{R}^{\ell}$.

3.2.4. Proposition .

Deux représentations d'un même système statistique, sont isomorphes .

Dém. si u et v sont deux représentations d'un même S.S., alors, on sait que $\quad \exists \lambda \in \mathbb{R}$ tel que

$$\forall i \in J_n, \forall j \in J_{\ell} , \; u'^{i}_{j} = \lambda v'^{i}_{j}$$

d'où $\quad < u'(e_j) | f_i > = \lambda < v'(e_j) | f_i >$

et $\quad u = \lambda v$.

Remarque : Nous allons montrer que la réciproque est fausse dans le cas général . Nous allons introduire une condition de " non-dégénérescence " d'un S.S., pour que la réciproque soit vraie .

3.2.4. Théorème

Soit u *et* v *deux représentations de même dimension et* $u = \sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i$ *la décomposition canonique de* u (*1.3.2)*.

Si u *et* v *sont isomorphes , alors, la décomposition canonique de* v *correspond à la même décomposition des espaces E et F en somme directe que* u *, et il existe* $\lambda \in \mathbb{R}$ *tel que*

$$v = \lambda \sum_{i=1}^{r} \sqrt{\lambda_i}\, Q_i .$$

Dém. Par hypothèse, il existe un opérateur T de F (resp. S de E) tel que

$$TT' = T'T = \alpha^2 I ,$$

$$SS' = S'S = \beta^2 I,$$

$$u(f_i) = Sv(f_i)$$

et $\qquad u'(e_j) = T'v'(e_j) .$

D'où $\qquad u = Sv = vT .$

On voit aisément que

$$\ker u = \ker v , \quad \mathrm{val}\, u = \mathrm{val}\, v ,$$

$$T(\ker u) = \ker u , \quad S(\mathrm{val}(v)) = \mathrm{val}\, v$$

En outre $T((\ker v)^{\perp}) = (\ker v)^{\perp}$ car

$$\forall x \in \ker v , \quad \forall y \in (\ker v)^{\perp} , \ \exists x' \in \ker v \ \text{ tel que } x = Tx'$$

et $\qquad < x \mid Ty > = < T'Tx' \mid y > = \alpha^2 < x' \mid y > = 0 .$

On peut donc supposer que $\ker(v) = \{0\}$ et $\mathrm{val}\,(v) = E$, c'est-à-dire que u et v sont bijectifs. Alors,

$$S = u \circ v^{-1}$$

et

$$T = v^{-1} \circ u .$$

Avec les notations 1.3.2., supposons que

$$u = \sum_{i=1}^{r} \sqrt{\lambda_i} . P_i .$$

Prenons $x = \sum_{i=1}^{r} x_i \in F = \bigoplus_{i=1}^{r} F_i$ où $x_i \in F_i$.

On sait que

$$\alpha^2 \|x\|^2 = \| (v^{-1} \circ u)(x) \|^2 = \|v^{-1}(\sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i(x_i)) \|^2$$

(1) $\quad \alpha^2 \|x_i\|^2 = \lambda_i \| v^{-1}(P_i(x_i))\|^2 = \alpha^2 \|P_i(x_i)\|^2 .$

Or $\qquad \|x\|^2 = \sum_{i=1}^{r} \|x_i\|^2$

d'où $\quad \| v^{-1}(\sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i(x_i))\|^2 = \sum_{i=1}^{r} \| v^{-1}\ (\sqrt{\lambda_i} P_i(x_i))\|^2$,

ou encore

$$\forall\, y = \sum_{i=1}^{r} y_i \in E = \bigoplus_{i=1}^{r} E_i \quad \text{où } y_i \in E_i ,$$

$$\|v^{-1}(\sum_{i=1}^{r} y_i)\|^2 = \sum_{i=1}^{r} \|v^{-1}(y_i)\|^2 .$$

Cela prouve que $\forall\, i \in J_r$, $\forall j \in J_r$, $i \neq j$ implique $v^{-1}(E_i) \perp v^{-1}(E_j)$.

En outre, d'après (1) , $\forall y_i \in E_i$,

$$\| v^{-1}(y_i)\| = \frac{\alpha}{\sqrt{\lambda_i}} \| y_i\| = \frac{1}{\sqrt{\mu_i}} \|y_i\|$$

où $\qquad \mu_i = \dfrac{\lambda_i}{\alpha^2}$.

En posant $\dfrac{1}{\sqrt{\mu_i}}\ v^{-1}|_{E_i} = Q'_i|_{E_i}$ et $Q'_i|_{E_i^{\perp}} = 0$,

nous avons la décomposition spectrale v

$$v = \sum_{i=1}^{r} \mu_i Q_i$$

avec $E = \bigoplus_{i=1}^{r} E_i$ et $F = \bigoplus_{i=1}^{r} v^{-1}(E_i)$. Pour démontrer que $v^{-1}(E_i) = F_i$ il faut recommencer un raisonnement analogue avec $S = u \circ v^{-1}$, au lieu de $T = v^{-1} \circ u$.

3.2.5. Corollaire

Si $\forall i \in J_r$, *dim* $F_i = 1$, *alors* u *et* v *isomorphes implique* $u = \lambda v$.

Dans ce cas u et v représentent le même S.S.. C'est la réciproque de 3.2.4. sous la condition $\dim F_i = 1$, $\forall i \in J_r$.

3.3. Système statistique non dégénéré .

3.3.1. Proposition .

Avec les notations de 1.3.2. , supposons que u soit une *représentation d'un S.S., dont la décomposition canonique soit*

$$u = \sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i \;,\; F = \Big(\bigoplus_{i=1}^{r} F_i\Big) \oplus \ker u \;,$$

$$E = \Big(\bigoplus_{i=1}^{r} E_i\Big) \oplus \ker u' \;.$$

S'il existe $i_0 \in J_r$ tel que $\dim F_{i_0} \geq 2$ alors il existe au moins une représentation v d'un système statistique différent, avec v isomorphe à u.

Dém : Considérons une base de $\bigoplus_{i \in J_r} F_i$ construite de la façon suivante :

$\forall i \in J_r$, soit $(\varphi^i{}_k)_k$ une base orthonormale de F_i.

Posons $\varepsilon^i{}_k = P_i(\varphi^i{}_k)$;

$(\varepsilon^i_k)_k$ est une base orthonormale de E_i. Soit v l'opérateur

$$v = \sum_{i \neq i_0} \sqrt{\lambda_i}\, P_i + \sqrt{\lambda_{i_0}}\; Q_{i_0}$$

où $Q_{i_0}(\varphi_1^{i_0}) = \varepsilon_2^{i_0}$, $Q_{i_0}(\varphi_2^{i_0}) = \varepsilon_1^{i_0}$ et $Q_{i_0}(\varphi_k^{i}) = \varepsilon_k^{i_0}$ pour $k \neq 1$ et $k \neq 2$.

D'après 3.2.4., v est isomorphe à u. Si v représentait le même S.S., alors on aurait : $\exists \lambda \in \mathbb{R}$ tel que $u = \lambda v$, ce qui est manifestement faux.

3.3.2. Définition

Nous dirons qu'un S.S. est " non-dégénéré ", si et seulement si, il admet au moins une représentation, dont l'opérateur d'inertie (3.1.6) a tous ses espaces propres de dimension 1 (i.e. $\forall i \in J_r$, $\dim F_i = 1$).

Tout S.S. ne vérifiant pas cette condition sera appelé "dégénéré".

3.3.3. Exemple : S.S. complètement dégénéré.

C'est le S.S. dont la représentation u est une isométrie de $\mathbb{R}^n$. La décomposition canonique de u est u lui-même. Tout autre isométrie de $\mathbb{R}^n$, différente de u et $-u$, représente nécessairement un S.S. différent.

3.3.4. Conclusion

De 3.3.1. nous déduisons que l'association S.S.-représentation, n'a un intérêt que pour les S.S. non-dégénérés . Ce sont les seuls que nous prendrons en considération dans la suite .

3.4. Analyse en composantes principales (ACP)

3.4.1. Théorème

Pour toute représentation u *d'un S.S. non-dégénéré , il existe une base orthonormée de val* u , $(\varepsilon_i)_{i \in J_r}$, *une base orthonormée de* (ker u)$^\perp$, $(\psi_i)_{i \in J_r}$ *et* r *nombres strictement positifs* $\lambda_1 > \lambda_2 > ..... > \lambda_r > 0$ *tel que*

$$u_p = \sum_{i=1}^{r} \sqrt{\lambda_i}\ \varepsilon_i \otimes \psi_i$$

où $\varepsilon_i \otimes \Psi_i : F \to E$ *et* $\forall x \in F , (\varepsilon_i \otimes \psi_i)(x) = < \psi_i \mid x > \varepsilon_i$.

Les ε_i *sont appelés " vecteurs principaux ", les* ψ_i *" composantes principales"* .

Ce théorème est un cas particulier de 1.3.2. Les ε_i sont les vecteurs propres de $V_p = u_p \circ u'_p$ à la valeur propre λ_i. Les ψ_i sont ceux de $U_p = u'_p \circ u_p$.

3.4.2. Interprétation

Pour décrire l'ensemble des individus$(u_p(f_i))_{i \in Jn}$, on les projettent sur les plans principaux $[\varepsilon_1,\varepsilon_2]$, $[\varepsilon_2,\varepsilon_3]$. La qualité de la projection d'un vecteur est donnée par le rapport

$$\|P_{[\varepsilon_1,\varepsilon_2]}(u_p(f_i))\| \,/\, \|u_p(f_i)\| .$$

La qualité globale est donnée par la part d'inertie expliquée

$$(\lambda_1 + \lambda_2)/\sum_{i=1}^{r} \lambda_i .$$

La même description est faite pour les paramètres $(u'_p(e_j))_{j \in J\ell}$, à l'aide des composantes principales .

Il est possible de représenter ces projections (des individus et des paramètres), sur les mêmes plans . Dans ce cas, l'axe " i " représentera ε_i et Ψ_i .

4. ANALYSE CANONIQUE (A.C)

4.1. Réduction d'une représentation

4.1.1. Définition

Soit deux représentations u_1 et u_2 *de deux S.S.. décrivant le même ensemble d'individus, de dimension* (n, ℓ_1) *et* (n, ℓ_2) *respectivement . On appellera " somme de* u_1 et u_2" *la représentation* $u = u_1 + u_2 : F \rightarrow E_1 \oplus E_2$ *de dimension* $(n, \ell_1 + \ell_2)$.

Le schéma est alors le suivant :

$$\begin{array}{ccc} E_1 \oplus E_2 & \xleftarrow{u_p} & F \\ \uparrow V_p & & \uparrow I \\ E1 \oplus E2 & \xrightarrow{u'_p} & F \end{array}$$

Si $V_{1,p} = u_{1,p} \circ u'_{1,p} = u_1 \circ D_p \circ u'_1$

$V_{2,p} = u_{2,p} \circ u'_{2,p} = u_2 \circ D_p \circ u'_2$

$V_{1,2} = u_1 \circ D_p \circ u'_2$

et $\quad V_{2,1} = u_2 \circ D_p \circ u'_1 = V'_{1,2}$,

alors,

$$V_p = u_p \circ u'_p = V_{1,p} + V_{2,p} + V_{1,2} + V_{2,1}$$

et $\quad U_p = u'_p \circ u_p = U_{1,p} + U_{2,p}$

Avec les notations matricielles

$$V_p = \left(\begin{array}{c|c} V_{1,p} & V_{1,2} \\ \hline V_{2,1} & V_{2,p} \end{array} \right)$$

Remarque.

Cette définition peut se généraliser à la somme d'un nombre fini de représentations. En particulier toute représentation de dimension (n, ℓ) peut-être considérée comme la somme de ℓ représentations de dimension $(n, 1)$. En effet, si

$$E_i = [\, e_i \,] \ , \ u_i = P_{E_i} \circ u \ ,$$

on voit que

$$E = \bigoplus_{i=1}^{\ell} E_i \ , \quad u = \sum_{i=1}^{\ell} u_i$$

$$u'_i\,(e_i) = u'\,(e_i)$$

et $\quad u'_i\,(e_j) = 0 \quad$ si $i \neq j$.

4.1.2. Définition

Avec les hypothèses de 4.1.1., nous dirons que u_1 *et* u_2 *sont "disjointes", quand* $u'_{1,p}\,(E_1) \perp u'_{2,p}\,(E_2)$. *Dans ce cas la somme sera dite "directe" et notée* $u = u_1 \oplus u_2$.

On voit que $V_p = \left(\begin{array}{c|c} V_{1,p} & 0 \\ \hline 0 & V_{2,p} \end{array} \right)$

et les sous-espaces propres de U_p sont, soit dans $u'_{1,p}(E_1)$, soit dans $u'_{2,p}(E_2)$:

$$u'_p(E_1 \oplus E_2) = u'_{1,p}(E_1) \oplus u'_{2,p}(E_2) .$$

Si $I_o(u)$ (resp. $I_o(u_1)$, $I_o(u_2)$) est l'indice d'inertie par rapport à O, de la représentation u (resp. u_1, u_2) , alors,

$$I_o(u) = I_o(u_1) + I_o(u_2) .$$

et une A.C.P. de u, sera moins précise que si on fait une A.C.P. de u_1 et de u_2 séparément .

Ce résultat peut se généraliser à un ensemble fini de représentations disjointes deux à deux .

4.1.3. Définition

Une représentation sera dite " réductible ", s'il existe une partition de l'ensemble des paramètres en deux sous-ensembles, dont tout élément de l' un est orthogonal à tout élément de l'autre.

Dans le cas contraire on dira que la représentation est " irréductible ".

4.1.4. Proposition

Toute représentation réductible peut s'écrire, comme somme directe de représentations.

Dém . supposons que u est une représentation de dimension (n , ħ) , que $J_\ell = M_1 \cup M_2$, où $M_1 \neq \emptyset$, $M_2 \neq \emptyset$ et $M_1 \cap M_2 = \emptyset$,

et que

$$\forall j \in M_1, \quad \forall j' \in M_2 \ , \ < u'_p(e_j) \,|\, u'_p(e_{j'}) > = 0 .$$

Notons $\quad F_i = [\{u'_p(e_j) \mid j \in M_i\}] \quad i = 1,2$;

$$F_1 \perp F_2 .$$

Soit u_1 et u_2 les applications définies par

$$u_i = u \circ p_{F_i} .$$

Si $E_i = u_i(F)$, alors $E_1 \perp E_2$ car $\forall x_1 \in E_1$, $\exists y_1 \in F_1$ tel que $x_1 = u_p(y_1)$ et $\forall x_2 \in E_2$, $< x_1 | x_2 > = < u_p(y_1) | x_2 > = < y_1 | u'_p(x_2) >$.

Or $u'_p(x_2) \in F_2$, puisque $u'_p(x_2)$ est une combinaison linéaire des $u'_p(e_j)$ avec $j \in M_2$; donc

$$< x_1 | x_2 > = 0$$

et

$$E = E_1 \oplus E_2 .$$

D'où $\quad u = u_1 \oplus u_2$.

4.1.5. Proposition

Toute représentation réductible, peut se décomposer en la somme directe fini de représentations irréductibles (disjointes deux-à-deux).

Dém. si u est de dimension (n, ℓ), il faut trouver une partition de $J_\ell = M_1 \cup M_2 \cup \ldots \cup M_s$ telle que les $F_i = [\{u'_p(e_j) \mid j \in M_i\}]$ soit de dimension la plus petite possible et orthogonaux deux-à-deux.

Pour trouver M_1 on procède de la façon suivante : on met dans M_1 l'indice 1 et tous les indices $j \in J_\ell$ tels que

$$< u'_p(e_1) | u'_p(e_j) > \neq 0 ;$$

puis pour chaque indice retenu on recommence comme pour l'indice 1.

J_ℓ étant fini, l'itération de ce procédé donnera M_1.

La construction des autres éléments de la partition est évidente.

Remarque : plutôt que de faire une A.C.P. d'une représentation réductible, on aura une information plus fine en la décomposant en somme directe de sous-représentations irréductibles et en faisant l'A.C.P. de chacune d'elle .

4.2. Analyse canonique de deux s.e.v.

4.2.1. Généralités

Soit F_1 et F_2 , deux s.e.v. de F, P_{F_1} et P_{F_2} les projecteurs orthogonaux correspondants .

Posons $A_1 = P_{F_1} P_{F_2} P_{F_1}$ et $A_2 = P_{F_2} P_{F_1} P_{F_2}$.

Comme $(P_{F_i})' = P_{F_i}$ et $\| P_{F_i} \| = 1$, on a $A_i \geqslant 0$ et $\| A_i \| \leqslant 1$, d'où spect $(A_i) \subset [\,0,1\,]$.

4.2.2. Théorème . Définition de l'A.C. de (F_1,F_2)

Avec les notations de 4.2.1. , il existe r *sous-espaces vectoriels* H_1^i (resp.H_2^i) de F_1 (resp.F_2) et r *nombres réels*

$0 < \beta_r < \ldots < \beta_2 < \beta_1 < 1$ *tels que*

$$F_1 = F_1 \cap F_2 \oplus \Big(\overset{r}{\underset{i=1}{\oplus}} H_1^i \Big) \oplus W_1 \cap W_2^{\perp} ,$$

$$F_2 = F_1 \cap F_2 \oplus \Big(\overset{r}{\underset{i=1}{\oplus}} H_2^i \Big) \oplus W_2 \cap W_1^{\perp} ,$$

$$A_1 = P_{F_1 \cap F_2} + \sum_{i=1}^{r} \beta_i P_{H_1^i} ,$$

$$A_2 = P_{F_1 \cap F_2} + \sum_{i=1}^{r} \beta_i P_{H_2^i} ,$$

$$\forall i \in J_r \ , \dim H_1^i = \dim H_2^i = n_i$$

Les sommes directes signifient que les s.e.v. sont orthogonaux.

Dém. nous savons , puisque A_i , $i = 1,2$, sont positifs , que

$$A_i = \sum_j \alpha_i^j P_{H_i^j}$$

où H_i^j est le sous espace propre de A_i à la valeur propre α_i^j .

En outre $\alpha_i^j \in [0,1]$. Procédons en plusieurs étapes :

(i) $F_1 \cap F_2$ est le sous espace propre de A_i, à la valeur propre 1. En effet

$$x \in F_1 \cap F_2 \Rightarrow A_i x = x,$$

et la réciproque se déduit du fait que

$$\|P_{F_i}(x)\| = \|x\| \Leftrightarrow x \in F_i .$$

(ii) $\ker A_i \supset F_i^{\perp}$.

(iii) $F_1 \cap F_2^{\perp} = F_1 \cap \ker A_1$. En effet l'inclusion $F_1 \cap F_2^{\perp} \subset F_1 \cap \ker A_1$ est évidente. Démontrons l'inclusion inverse,

soit $x \in F_1 \cap \ker A_1$, $x = y+z$ avec $y \in F_2$ et $z \in F_2^{\perp}$, d'où $\|x\|^2 = \|y\|^2 + \|z\|^2$.

$$A_1(x) = P_{F_1}(P_{F_2}(x)) = P_{F_1}(y) = 0 ;$$

donc $y \in F_1^{\perp}$ et $z = x-y$ avec $x \in F_1$ et $y \in F_1^{\perp}$, d'où $\|z\|^2 = \|x\|^2 + \|y\|^2$.

En comparant avec la précédente égalité entre les normes, on tire $\|y\| = 0$, d'où $y = 0$ et $x \in F_1 \cap F_2^{\perp}$..

(IV) $\ker(A_1) = F_1 \cap F_2^{\perp} \oplus F_1^{\perp}$. L'inclusion $F_1 \cap F_2^{\perp} \oplus F_1^{\perp} \subset \ker A_1$ est évidente. Montrons l'inclusion inverse,

soit $x \in \ker A_1$, $x = y+z$ avec $y \in F_1$ et $z \in F_1^{\perp}$;

d'après (ii) $y = x-g \in \ker A_1$ et d'après (iii) $y \in F_1 \cap F_2^{\perp}$.

On vient donc de montrer que $F_1 \cap F_2^{\perp} \oplus F_1^{\perp}$ est le sous espace propre de A_1 à la valeur propre 0. On montrera de la même façon que

$$\ker A_2 = F_2 \cap F_1^{\perp} \oplus F_2^{\perp} .$$

(V) De ce qui précède on déduit que

$$A_1 = P_{F_1 \cap F_2} + \sum_{i=1}^{r} \beta_i P_{H_1^i}$$

avec $0 < \beta_r < \dots < \beta_2 < \beta_1 \leq 1$. Montrons que

$$A_2 = P_{F_1 \cap F_2} + \sum_{i=1}^{r} \beta_i \, P_{H_2^i}$$

où H_2^i est l'espace propre de A_2 à la valeur propre β_i et que $\dim H_2^i = \dim H_1^i$.

$\forall x \in H_1^i$, $A_2(P_{F_2}(x)) = P_{F_2} P_{F_1} P_{F_2} P_{F_1} x = P_{F_2}(A_1(x)) = \beta_i P_{F_2}(x)$.

D'où β_i est une valeur propre de A_2, $P_{F_2}(H_1^i) \subset H_2^i$ et $\dim H_1^i \leq \dim H_2^i$

puisque $P_{F_2}\big|_{H_1^i}$ est injectif. L'inclusion et l'inégalité inverse se

démontrent de la même façon.

(Vi) H_1^i est orthogonal à $\ker A_1$ puisque ce sont deux sous espaces propres correspondant à des valeurs propres différentes. De (iV), on déduit que $H_1^i \subset F_i$ et de (iV) et (V) on déduit que

$$F_1 = F_1 \cap F_2 \oplus \left(\bigoplus_{i=1}^{r} H_1^i\right) \oplus F_1 \cap F_2^{\perp}$$

et de la même façon

$$F_2 = F_1 \cap F_2 \oplus \left(\bigoplus_{i=1}^{r} H_2^i\right) \oplus F_2 \cap F_1^{\perp}$$

4.2.3 <u>Proposition</u>

$$F_1 + F_2 = F_1 \cap F_2 \oplus \left(\bigoplus_{i=1}^{r} (H_1^i + H_2^i)\right) \oplus F_1 \cap F_2^{\perp} \oplus F_2 \cap F_1^{\perp}.$$

<u>Dém.</u> il est évident que H_1^i est orthogonal à $F_1 \cap F_2$, à $F_1 \cap F_2^{\perp}$

et à $F_2 \cap F_1^{\perp}$. Il en est de même de H_2^i. Il reste à démontrer que $H_1^i + H_2^i$

est orthogonal à $H_1^j + H_2^j$ si $i \neq j$, ou ce qui revient au même que H_1^i est

orthogonal à H_2^j. En effet, si $x \in H^i$ et $y \in H^j$, alors,

$$< x|y > = < x|\, P_{F_2} y > = < P_{f_2} x|\, y > = 0$$

puisque $P_{F_2} x \in H_2^i$ et que H_2^i est orthogonal à H_2^j.

4.2.3. Interprétation géométrique

4.2.3.1. Notons

$$L_1 = F_1 - F_1 \cap F_2 = \Big(\bigoplus_{i=1}^{r} H_1^i\Big) \oplus F_1 \cap F_2^{\perp}$$

$$L_2 = F_2 - F_1 \cap F_2 = \Big(\bigoplus_{i=1}^{r} H_2^i\Big) \oplus F_2 \cap F_1^{\perp}$$

Nous allons chercher les éléments normés de L_1 les plus proches de F_2.

Tout élément normé de $F_1 \cap F_2^{\perp}$ étant orthogonal à L_2, sa distance à L_2 est égale à 1. Si $x \in H_1^i$, alors, $x \perp F_2 \cap F_1^{\perp}$ et $x \perp H_2^j$ pour $j \neq i$.

Donc $\quad d(x,F_2) = d(x, H_2^i)$. Or

$$d(x,F_2) = \| x - P_{F_2}(x) \|$$

$$\begin{aligned} d(x,F_2)^2 &= \langle x - P_{F_2}(x) | x - P_{F_2}(x) \rangle = \\ &= \|x\|^2 - 2 \langle x | P_{F_2}(x) \rangle + \langle P_{F_2}(x) | P_{F_2}(x) \rangle \\ &= \|x\|^2 - \langle x | P_{F_2}(x) \rangle = 1 - \langle P_{F_1}(x) | P_{F_2}(x) \rangle \\ &= 1 - \langle x | A_1 x \rangle = 1 - \beta_i \text{, d'où} \end{aligned}$$

Proposition

$\forall x \in H_1^i$, si $\|x\| = 1$, *alors*

$$d(x,F_2) = \sqrt{1-\beta_i}$$

et le vecteur normé de F_2 *le plus proche de* x

est $P_{F_2}(x) / \|P_{F_2}(x)\| = P_{F_2}(x) / \sqrt{\beta_i} \in H_2^i$.

4.2.3.2. Proposition

$\forall x \in L_1$, si $\|x\| = 1$, *alors*,

$$d(x,F_2) \geqslant \sqrt{1-\beta_1}.$$

Dém. autrement dit, les vecteurs normés de L_1 les plus proches de F_2 sont ceux de H_1^1. Si $x = \sum_{i=1}^{r} x_i + y$, avec $x_i \in H_1^i$, $y \in F_1 \cap F_2^{\perp}$

et $\|x\| = \sum_{i=1}^{r} \|x_1\|^2 + \|y\|^2 = 1$, alors ,

$$d(x,F_2)^2 = < x - P_{F_2}(x) \mid x - P_{F_2}(x) >$$

$$= 1 - \sum_{i=1}^{r} < x_i \mid P_{F_2}(x_i) >$$

$$= 1 - \sum_{i=1}^{r} \beta_i \|x_i\|^2 \geqslant 1 - \beta_i$$

puisque $\beta_1 \geqslant \sum_{i=1}^{r} \beta_i \|x_i|^2$.

4.2.3.3. Interprétation

La décomposition de F_1 que nous avons obtenu, regroupe ses vecteurs en sous espaces vectoriels de plus en plus éloignés de F_2.

En effet,

$$F_1 = F_1 \cap F_2 \oplus \left(\bigoplus_{i=1}^{r} H_1^i \right) \oplus F_1 \cap F_2^{\perp}$$

et $\forall x \in F_1$, $\|x\| = 1$, $x \in F_1 \cap F_2 \Rightarrow d(x,F_2) = 0$

$$x \in H_i^i \Rightarrow d(x,F_2) = \sqrt{1-\beta_i} < 1$$

$$x \in F_1 \cap F_2^{\perp} \Rightarrow d(x,F_2) = 1 .$$

4.3. Analyse canonique de deux représentations

4.3.1. Soit u_1 et u_2 , deux représentations de deux S.S. sur un même ensemble d'individus . On a le schéma suivant:

$$\begin{tikzcd}
E_1 & F \arrow[l, "u_1"'] \arrow[r, "u_2"] & E_2 \\
E_1 \arrow[u, "V_1"] \arrow[r, "u'_1"] & F \arrow[u, "I"] & E_2 \arrow[u, "V_2"] \arrow[l, "u'_2"']
\end{tikzcd}$$

On pose $F_1 = u'_1(E_1)$, $F_2 = u'_2(E_2)$. F_1 (resp.F_2) s'appellera " espace expliqué " par u_1(resp.u_2) . En effet, tout élément de F_1(resp.F_2) est combinaison linéaire des paramètres de u_1 (resp.u_2) .

Définition

Nous dirons que u_1 est " moins fin " que u_2 , ou ce qui revient au même , u_2 est " plus fin " que u_1 si et seulement si,

$$F_1 \subset F_2 .$$

Lorsqu'il y aura égalité , nous dirons que u_1 et u_2 sont "équivalents " .

Nous avons là, une relation d'ordre partiel , sur l'ensemble des représentations de tous les S.S. sur un même ensemble d'individus .

4.3.2. Définition

Quand u_1 et u_2 ne sont pas comparables (4.3.1), on appellera analyse canonique (A.C.) de(u_1, u_2), l'analyse canonique de(F_1,F_2).

Cela permet de comparer de façon précise les espaces expliqués par chacune des représentations . En effet :

$$F_1 = F_1 \cap F_2 \oplus (\bigoplus_{i=1}^{r} H_1^i) \oplus F_1 \cap F_2^{\perp}$$

$$F_2 = F_1 \cap F_2 \oplus (\bigoplus_{i=1}^{r} H_2^i) \oplus F_2 \cap F_1^{\perp}$$

* $F_1 \cap F_2$ est la partie de F_1 (resp.F_2) "totalement expliquée " par u_2(resp.u_1) .

* H_1^i (resp. H_2^i) est " partiellement expliquée " par u_2 (resp.u_1) et ce sont les vecteurs de H_2^i (resp.H_1^i) qui " l'expliquent " le mieux, car les éléments de F2 (resp.F1) les plus proches de H_1^i (resp.H_2^i)sont ceux de H_2^i (resp.H_1^i) et leur distance à H_1^i (resp.H_2^i) est égale à $\sqrt{1-\beta i}$ (4.2.3.1.).

* $F_1 \cap F_2^{\perp}$ (resp. $F_2 \cap F_1^{\perp}$) est la partie de F_1 (resp. F_2)

" non expliquée " par u_2 (resp.u_1) .

4.3.3. <u>Définition</u>

Soit $f_1, \ldots, f_q$ *une base orthonormée de* $F_1 \cap F_2$, $f_{i,1}, \ldots, f_{i,n_i}$ *une base orthonormée de* H_1^i . *Nous savons que* $g_{i,1}, \ldots, g_{i,n_i}$ *où*

$$g_{i,k} = \frac{P_{F_2}(f_{i,k})}{\sqrt{\beta_i}}$$

est une base orthonormée de H_2^i . *Les couples* $(f_{i,k}, g_{i,k})$ *sont les " couples de paramètres canoniques "* .

En projetant sur les plans construits à l'aide des $(f_i)_i$ et des $(f_{i,k})_{i,k}$ les paramètres des deux représentations, on peut avoir une idée de leur " proximité ". La " fidélité " de la projection d'un paramètre est d'autant meilleure que le rapport , norme de la projection/ norme du paramètre projeté , est proche de 1 .

4.3.4. <u>Facteurs canoniques, vecteurs canoniques .</u>

Considérons le diagramme suivant

$$\begin{tikzcd}
E_1 \arrow[d, "V_1^{-1}"'] & F \arrow[l, "u_1"'] & E_2 \arrow[l, "u'_2"'] \arrow[dll, shift right, "B_1"'] \\
E_1 \arrow[r, "u'_1"'] \arrow[urr, shift right, "B_2"'] & F \arrow[r, "u_2"'] & E_2 \arrow[u, "V_2^{-1}"']
\end{tikzcd}$$

Pour la définition de V_1^{-1} et V_2^{-1} voir 1.2.9. u'_1 (resp. u'_2) étant injectif sur val(u_1) (resp. val(u_2)), il existe un " facteur canonique " et un seul :

* $a^1_{i,k} \in$ val(u_1) (resp.$a^2_{i,k} \in$ val(u_2) tel que

$u'_1(a^1_{i,k}) = f_{i,k}$ (resp. $u'_2(a^2_{i,k}) = g_{i,k}$)

* $a^1_k \in val(u_1)$ (resp. $a^2_k \in val(u_2)$)

tel que

$u'_1(a^1_k) = f_k$ (resp. $u'_2(a^2_k) = f_k$).

On appelle " vecteurs canoniques " les éléments

$b'_{i,k} = u_1(f_{i,k})$ $\quad b^2_{i,k} = u_2(g_{i,k})$

$b^1_k = u_1(f_k)$ $\quad b^2_k = u_2(f_k)$.

On voit facilement que

$u_1(g_{i,k}) = \sqrt{\beta_i}\, b^1_{i,k}$, $u_2(f_{i,k}) = \sqrt{\beta_i}\, b^2_{i,k}$,

$V_1(a^1_{i,k}) = b^1_{i,k}$ $\quad V_1(a^1_k) = b^1_k$

$V_2(a^2_{i,k}) = b^2_{i,k}$ $\quad V_2(a^2_k) = b^2_k$

Posons $B_1 = \overset{-1}{V_1} \circ u_1 \circ u'_2$ et $B_2 = \overset{-1}{V_2} \circ u_2 \circ u'_1$.

On voit facilement que $(B_1 \circ B_2)(a^1_{i,k}) = \beta_i\, a^1_{i,k}$

$(B_1 \circ B_2)\ (a^1_k) = a^1_k$

$(B_2 \circ B_1)\ (a^2_{i,k}) = \beta_i\, a^2_{i,k}$

$(B_2 \circ B_1)\ (a^2_k) = a^2_k$

4.3.5. <u>Remarque</u> : A.C.P.

L'étude de la proximité des paramètres de u_1 et de u_2 peut se faire aussi, par l'A.C.P. de $u_1 + u_2$.

La projection des paramètres sur les plans principaux, donnera une description des positions relatives, qui ne se superpose pas à celle donnée par l'A.C.. En effet, elle tient compte, pour définir ces plans, de l'indice d'inertie de chaque représentation.

La fidélité de la description des $(u'_{1,p}(e_j))_{j\in J\ell_1}$ (resp. $(u'_{2,p}(h_j))_{j\in J\ell_2}$ dans cette projection est donnée par la "part d'inertie expliquée", c'est-à-dire, le rapport de l'indice d'inertie des projections des $u'_{1,p}(e_j)$(resp. $u'_{2,p}(h_j)$) sur $I_0(u_1)$(resp. $I_0(u_2)$).

La fidélité de la description de chaque paramètre, est donnée par le rapport de la norme de la projection sur celle du paramètre projeté.

4.3.6. Remarque importante : paramètres non centrés.

Si u_p est une représentation dont les paramètres ne sont pas tous centrés, alors, la représentation v_p correspondant aux paramètres centrés est

$$v_p = u_p \circ P_{[\varphi_p]^{\perp}} \qquad (3.1.3.3.)$$

Proposition

u_p et v_p *ne sont jamais équivalents. Plus précisément, si* $\varphi_p \in u'_p(E)$ *alors* u_p *est strictement plus fine que* v_p *et si* $\varphi_p \notin u'_p(E)$ *alors dim* $u'_p(E) = \dim v'_p(E)$ et u_p *n'est pas comparable à* v_p.

Dém.

si $\varphi_p \in u'_p(E)$, il est évident que $\dim v'_p(E) = \dim u'_p(E) - 1$,

et que $v'_p(E) \subset u'_p(E)$.

si $\varphi_p \notin u'_p(E)$, alors la restriction de $P_{[\varphi_p]^{\perp}}$ à $u'_p(E)$ est injective.

D'où $\dim(u'_p(E)) = \dim(v'_p(E))$

alors que $u'_p(E) \neq v'_p(E)$.

On comprend pourquoi, nous n'avons pas remplacé systématiquement dans ce qui précède, u_p par v_p, d'autant plus que cela ne simplifie pas l'exposé . En outre, comme nous le verrons dans les chapitres suivants, les " modalités " d'un paramètre quantitatif ou qualitatif ne sont jamais centrés. On pourra donc leur appliquer , sans rien changer, tous les résultats qui précèdent .

4.4. Changement d'unité des paramètres .

4.4.1. Faire un changement d'unité pour le paramètre $u'_p(e_j)$, revient à le multiplier par un scalaire $\lambda_j > 0$.

Définition

Un changement d'unité des paramètres d'une représentation u, consiste à remplacer u par $\delta_\lambda \circ u$ où

$$\lambda = (\lambda_1,\lambda_2,\dots \lambda_\ell) \in (\mathbb{R}^{+*})^\ell \quad \text{et} \quad \delta_\lambda = \sum_{i=1}^{\ell} \lambda_i P_{[e_i]} .$$

Il est évident que δ_λ est une bijection linéaire de E .

Notons $(\delta_\lambda)^{-1} = \delta_{1/\lambda}$.

4.4.2. Si l'on choisit pour λ_j , la norme $\|u'_p(e_j)\|$ (resp.l'écart-type σ_j , si les paramètres sont centrés), alors, la représentation $\delta_{1/\lambda} \circ u$ (resp. $\delta_{1/\sigma} \circ u$)correspond aux paramètres normés(resp.réduits) et la matrice d'inertie (resp.variance-covariance) est

$$\delta_{1/\lambda} \circ V_p \circ \delta_{1/\lambda} (\text{resp. } \delta_{1/\sigma} \circ V_p \circ \delta_{1/\sigma}$$

qui est la matrice des corrélations) .

L'intérêt de ce changement d'unité est que les paramètres sont sur la sphère unité . Leur projection sur les plans principaux sera d'autant plus fidèle qu'elle sera proche du cercle unité .

4.4.3. Proposition

Les représentations u et $\delta_\lambda \circ u$ *sont isomorphes si et seulement si*

$$\lambda_1 = \lambda_2 = \dots = \lambda_\ell .$$

Dém . évidente à cause de 3.2.5.

4.4.4. Proposition

Les représentations u *et* δ_λ o u *sont équivalentes.*

Dém. évidente puisque δ_λ est une bijection .

5. PARAMETRES QUALITATIFS
ANALYSE FACTORIELLE DISCRIMINANTE (A.F.D.)

5.1. Paramètres qualitatifs

5.1.1. Définition

Soit u , *une représentation d'un S.S. de dimension* (n,ℓ) . *Pour tout* $M \subset J_n$, *posons*

$$\varphi_M = \sum_{i \in M} \sqrt{p_i}\, f_i$$

φ_M *s'appellera une " modalité ".*

Il est clair que $\quad \| \varphi_M \|^2 = \sum_{i \in M} p_i = P(M)$,

$P(M)$ est appelé la " probabilité de la modalité φ_M " . La moyenne de φ_M est $< \varphi_M | \varphi_p > = P(M)$ qui n'est donc jamais nul, et sa variance est

$$< \varphi_M | P_{[\varphi_p]^\perp} \varphi_M > = P(M) - P(M)^2 = P(M)\,(1-P(M)) .$$

5.1.2. Tout paramètre $x \in F$, définit une partition de $J_n = M^1 \cup M^2 \cup \ldots\ldots \cup M^r$ par

$\forall i \in M^k \; , \; \forall i' \in M^{k'} \quad , < x | f_i > = < x | f_{i'} > \Longleftrightarrow k = k'$.

Si nous notons x_{M^k} , la composante de x sur les $(f_i)_{i \in M^k}$, alors

$$x_p = D_p(x) = \sum_{k \in J_r} x_{M^k}\, \varphi_{M^k}$$

les $(\varphi_{M^k})_{k \in J_r}$ sont les " modalités " de x_p .

Elles sont orthogonales deux à deux .

La fonction de probabilité f de x est définie par

$$f(\alpha) = 0 \text{ si } \alpha \notin \{ x_{M^1}, \ldots\ldots x_{M^r} \}$$

et

$$f(x_{M^k}) = P(M^k)$$

Définition

On appelle " potentiel de prévision " de x, la s.e.v. F_x *engendré par ses modalités* .

Autrement dit, F_x est l'ensemble de tous les paramètres $y_p = D_p(y)$ ayant les mêmes modalités que x(ou x_p) .

5.1.3. Définition

$\forall x \in F$, $\forall y \in F$, x *sera dit " moins fin " que* y, *ou* y *" plus fin " que* x , *si et seulement si* $F_x \subset F_y$.

Nous obtenons ainsi une relation d'ordre partiel sur l'ensemble des paramètres .

5.1.4. Définition

x *sera dit " équivalent " à* y *si et seulement si* $F_x = F_y$ *et les classes d'équivalence seront appelées " paramètre qualitatif ", de l'ensemble des individus* J_n .

Autrement dit, les paramètres qualitatifs entièrement définis par leurs modalités . Ils sont donc en bijection avec l'ensemble des partitions de J_n.

5.1.5. Partitions de J_n.

Nous dirons que la partition $(M^k)_{k \in J_{s_s}}$ est " plus fine " que la partition $(N^\ell)_{\ell \in J_t}$ et nous écrirons $(M^k)_{k \in J_s} \leqslant (N^\ell)_{\ell \in J_t}$, si et seulement si,

$\forall k \in J_S$, $\exists \ell \in J_t$ tel que $M^k \subset N^\ell$.

Donc tout N^{ℓ} est réunion de M^k.

Pour cette relation d'ordre partiel, l'ensemble des partitions forme un treillis, c'est-à-dire quelque soit deux partitions $(M^k)_{k\in J_s}$ et $(N^{\ell})_{\ell\in J_t}$, il existe un sup. noté $(M^k)_{k\in J_s} \vee (N^{\ell})_{\ell\in J_t}$ et un inf. noté $(M^k)_{k\in J_s} \wedge (N^{\ell})_{\ell\in J_t} = (M^k \cap N^{\ell})_{(k,\ell)\in J_s \times J_t}$.

5.1.6. Représentation d'un paramètre qualitatif.

Considérons le paramètre qualitatif correspondant à la partition $(M^k)_{k\in J_s}$ de J_n. Il peut être considéré comme un S.S. de dimension (n,s), les paramètres du système étant les fonctions caractéristiques des $(M^k)_{k\in J_s}$. Soit $(h_k)_{k\in J_s}$ la base canonique de $\mathbb{R}^s$. La représentation v du S.S. est définie par $v : F \to \mathbb{R}^s$

et $\forall k\in J_s$, $\forall i\in M^k$, $v(f_i) = h_k$.

On voit aisément que

$$V_p(h_k) = P(M^k)\, h_k \text{ où } V_p = v.\, D_p \circ v'$$

et

$$U_p(\varphi_{M^k}) = P(M^k)\, \varphi_{M^k} \text{ où } U_p = D_{\sqrt{p}} \circ v' \circ v \circ D_{\sqrt{p}}.$$

Donc les $(h_k)_{k\in J_s}$ sont les vecteurs principaux et les $(\varphi_{M^k})_{k\in J_s}$ les composantes principales. Le S.S. ainsi définit est non-dégénérée que si les $(P(M^k))_{k\in J_s}$ sont distincts deux-à-deux. C'est évidemment la condition pour que cette représentation ait un intérêt (3.3.4).

5.2. Analyse factorielle discriminante (A.F.D).

5.2.1. Considérons un S.S. de dimension $(n,\ell+1)$. Supposons que pour l'un des paramètres nous nous intéressions uniquement à ses modalités ou ce qui revient au même supposons qu'il y a ℓ paramètres quantitatifs et un qualitatif.

Pour étudier un tel système, nous prendrons la représentation u du S.S. de dimension (n,ℓ) comprenant les ℓ paramètres quantitatifs, et la représentation v du S.S. de dimension (n,s) associé au paramètre qualitatif.

5.2.2. Définition

On appelle A.F.D. l'analyse canonique du couple (u,v).

Mais comme nous l'avons déjà vu l'étude des proximités entre les paramètres de u et les modalités de v peut se faire aussi l'aide de l'A.C.P.. de $u+v$.

5.3. Classes d'individus définies par un paramètre .

5.3.1. Soit u une représentation d'un S.S. de dimension (n,ℓ) .

Supposons que nous nous intéressions à une seule modalité φ_M , $M \subset J_n$, d'un paramètre qualitatif quelconque . Notons

$$F_M = [\{ f_i | i \in M \}]$$

et

$$\varepsilon_M = \frac{1}{P(M)} \varphi_M = \frac{1}{P(M)} \sum_{i \in M} \sqrt{p_i}\, f_i$$

Nous cherchons à décrire d'une part, la classe A(M) des individus présentant la modalité φ_M(i.e. $A(M) = \{u_p(f_i) | i \in M\}$) et d'autre part, la proximité des paramètres $u'_p(e_j)$ avec la modalité φ_M.

Si nous posons $u_{p,M} = u_p \circ P_{F_M}$, on voit que

* $\forall i \in M$, $u_{p,M}(f_i) = u_p(f_i)$;
* $\forall i \notin M$, $u_{p,M}(f_i) = 0$;
* $\forall j \in J_\ell$ $u'_{p,M}(e_j) = P_{F_M}(u'_p(e_j))$ est la projection des paramètres sur l'espace F_M qui contient φ_M ;
* $u_{p,M}(\varepsilon_M) = g_M = \sum_{i \in M} \frac{p_i}{P(M)} u(f_i)$ est le centre de gravité de la classe A(M), plus précisément celui des $(u(f_i))_{i \in M}$.

Le problème posé sera donc résolu en faisant une A.C.P. de $u_{p,M}$ et en projetant φ_M sur les plans principaux des paramètres .

5.3.2. Centres de gravités " intra-classes " .

En reprenant les notations 5.2.. et 5.3.1. , les centres de gravité " intra-classe " sont les

$$g_{M^k} = u_{p,M^k}(\varepsilon_{M^k}) = \sum_{i \in M^k} \frac{p_i}{P(M^k)} u(f_i) .$$

Si chaque g_{M^k} est muni du poids $P(M^k)$, on voit que

$$g = \sum_{k \in J_s} P(M^k)\, g_{M^K} = \sum_{i \in J_n} p_i\, u(f_i)$$

est le centre de gravité des $(u(f_i))_{i \in J_n}$.

Pour décrire la répartition des $(g_k)_{k \in J_s}$ dans E, il suffit de prendre un espace euclidien G muni d'une base $\gamma_1, \ldots \gamma_s$ orthonormale , chaque γ_k étant muni du poids $P(M^k)$; et le problème sera résolu par l'A.C.P. de la représentation

$$\begin{aligned} w : G &\longrightarrow E \\ \gamma_k &\longrightarrow w(\gamma_k) = g_{M^k} . \end{aligned}$$

5.3.3. Qualité de la description des classes par leur centre de gravité .

Notons $r_k = \min_{\ell \in J_s} \| g_{M^k} - g_{M^\ell} \|$,

$B_k = B(g_{M^k}, r_k)$ boule ouverte de centre g_{M^k} et de rayon r_k ; c'est la plus grande boule possible centrée en g_{M^k} et ne contenant qu'un seul centre de gravité .

$$C_k = B\left(g_{M^k}, r_k/_2\right)$$

La famille des $(B_k)_{k\in J_s}$ est un recouvrement de l'ensemble Γ des centres de gravité par des boules ouvertes les plus grandes possibles, séparant Γ . Les $(B_k)_{k\in J_s}$ ne sont pas disjoints deux-à-deux . Par contre les $(C_k)_{k\in J_s}$ est un recouvrement du même ensemble par des boules ouvertes les plus grandes possibles, disjointes deux-à-deux .

La qualité de la représentation de la classe A_k par son centre de gravité g_{M^k} peut-être représentée par deux coefficients :

α_k(resp.β_k) = proportion des éléments de A_k appartenant à B_k (resp. C_k) .

Il va de soi que si α_k et β_k sont très faibles, la description de A_k par g_{M^k} n'a pas un grand intérêt . En tout cas , pas plus que la moyenne d'une variable aléatoire dont l'écart-type est très grand .

La qualité de la représentation globale des classes $(A_k)_{k\in J_s}$ par leurs centres de gravité $(g_{M^k})_{k\in J_s}$, peut-être représentée par deux coefficients :

α(resp. β) = somme sur k des nombres d'éléments de A_k appartenant à B_k (resp.C_k) , divisé par n . C'est donc la proportion des individus séparées par les $(B_k)_{k\in J_s}$ (resp.$(C_k)_{k\in J_s}$) .

6. INDICE DE CORRELATION
ANALYSE FACTORIELLE DES CORRESPONDANCES (A.F.C.)

6.1. Couple de paramètres irréductibles.

6.1.1 . Définition

Considérons deux paramètres x et y *quantitatifs ou qualitatifs sur un même ensemble d'individus* J_n . *Soit* $(M^k)_{k\in J_s}$ *et* $(N^\ell)_{\ell\in J_t}$ *les partitions respectives qu'ils induisent sur* J_n . *Nous dirons* *que* (x,y) *est " irréductible " si et seulement si*

$$(M^k)_{k\in J_s} \vee (N^\ell)_{\ell\in J_t} = J_n \quad (5.1.5.)\,.$$

Quand la relation ne sera pas vérifiée nous dirons que (x,y) *est " réductible "*.

Remarque : si u et v sont les représentations associées aux paramètres qualitatifs correspondants, alors ,(x, y) est irréductible si et seulement si u+v est irréductible(4.1.3) .

6.1.2. Proposition

(x,y) *est irréductible si et seulement si*

$\forall k\in J_s$, $\forall \ell\in J_t$, $< \varphi_{M^k} \mid \varphi_{N^\ell} > \neq 0$ *ou* $\exists\, k'\in J_s$, $\exists\, \ell'\in J_t$

tel que $< \varphi_{M^k} | \varphi_{N^{\ell'}} > \neq 0$, $< \varphi_{N^\ell} | \varphi_{M^{k'}} > \neq 0$ et $< \varphi_{M^{k'}} | \varphi_{N^{\ell'}} > \neq 0$.

Dém. on utilise le fait que les $(\varphi_{M^k})_{k\in J_s}$ (resp. $(\varphi_{N^\ell})_{\ell\in J_t}$ sont orthogonaux deux à deux et les définitions 4.1.3. et 6.1.1.

6.1.3. Corollaire

Si x *et* y *sont équivalents*(i.e. s=t et $\forall k\in J_s$, $M^k = N^k$), *alors*, (x,y) *est irréductible si et seulement si* x *et* y *n'ont qu'une seule modalité* φ_p.

6.1.4. Proposition

(x,y) *est irréductible, si et seulement si* , $F_x \cap F_y = [\,\varphi_p\,]$.

Dém. $\varphi_p = \sum_{k\in J_s} \varphi_{M^k} = \sum_{\ell\in J_t} \varphi_{N^\ell} \in F_x \cap F_y$.

Supposons que $\sum_{k\in J_s} \alpha_k \varphi_{M^k} = \sum_{\ell\in J_t} \beta_\ell \varphi_{N^\ell}$ et montrons que $\forall k\in J_s$, $\forall \ell\in J_t$, $\alpha_k = \beta_\ell = \gamma$.

En effet, chaque f_i n'intervient qu'une fois dans chaque membre.

Si $< \varphi_{M^k} \mid \varphi_{N^\ell} > \neq 0$, cela signifie que $\exists\, i \in J_n$ tel que $f_i \in M^k \cap N^\ell$ et ses coefficients sont égaux .

$$\alpha_k \sqrt{p_i} = \beta_\ell \sqrt{p_i} \qquad \text{d'où } \alpha_k = \beta_\ell \ .$$

Si $< \varphi_k \mid \psi_\ell > = 0$, on utilise 6.1.2. et le raisonnement ci-dessus .

6.2. Paramètres irréductibles et indice de corrélation .

6.2.1. Définition

On dit que deux modalités quelconques φ_M et φ_N *sont indépendantes si et seulement si*

$$P(M \cap N) = P(M).\, P(N).$$

Remarque : si $M \cap N = \emptyset$, alors $P(M \cap N) = 0$ et $P(M).P(N) \neq 0$. Cela implique que deux modalités d'un même paramètre ne peuvent pas être indépendantes .

6.2.2. Proposition

φ_M et φ_N *sont indépendantes , si et seulement si, leurs projections sur* $[\varphi_p]^\perp$ *sont orthogonales* .

Dém. $P(M \cap N) = < \varphi_M \mid \varphi_N >$

$P(M) = < \varphi_M \mid \varphi_p >$, $P(N\ ) = < \varphi_N \mid \varphi_p >$.

D'où $P(M).P(N) = < \varphi_M \mid \varphi_p > < \varphi_p \mid \varphi_N > = < \varphi_M \mid P_{[\varphi_p]} \varphi_N >$,

et $P(M \cap N) - P(M)\, P(N) = < \varphi_M \mid P_{[\varphi_p]^\perp} \varphi_N > = 0$.

6.2.3. Définition

Avec les notations 6.1.1. , nous dirons que x *et* y *sont indépendants si et seulement si* $\forall k \in J_s$, $\forall \ell \in J_t$, φ_{M^k} et φ_{N^ℓ}

sont indépendants (i.e. $P(M^k \cap N^\ell) = P(M^k).\,P(N^\ell)$) .

Remarque . cette définition , qui est la définition classique de l'indépendance de deux paramètres , implique que (x,y) est irréductible .

6.2.4. Lemme

$$\sum_{k\in J_s} \sum_{\ell\in J_t} \frac{P(M^k \cap N^\ell)^2}{P(M^k)\,P(N^\ell)} = \mathrm{Tr}\,(P_{F_x} \,.\, P_{F_y}) .$$

Dém posons $\psi_k = \varphi_{M^k}/\|\varphi_{M^k}\| = \varphi_{M^k}/\sqrt{P(M^k)}$

$$\chi_\ell = \varphi_{N^\ell}/\|\varphi_{N^\ell}\| = \varphi_{N^\ell}/\sqrt{P(N^\ell)}$$

Sachant que $(\psi_k)_{k\in J_s}$ (resp.$(\chi_\ell)_{\ell\in J_t}$) est une base orthonormale de F_x (resp.F_y) , on voit que :

$$\sum_{k\in J_s} \sum_{\ell\in J_t} \frac{P(M^k \cap N^\ell)^2}{P(M^k)P(N^\ell)} = \sum_{k,\ell} <\psi_k \mid \chi_\ell>^2$$

$$= \sum_k \|P_{F_y} \psi_k\|^2 = \sum_k <\psi_k \mid P_{F_y} \psi_k>$$

$$= \sum_k <\psi_k | P_{F_x} P_{F_y} \psi_k> = \mathrm{Tr}\,(P_{F_x} \,.\, P_{F_y}) .$$

Le calcul de cette trace se fait par l'A.C. de (F_x, F_y) .

6.2.5. Proposition

x et y sont indépendants , si et seulement si, $\mathrm{Tr}(P_{F_x} P_{F_y}) = 1$ *et dans ce cas*

$$F_x = [\varphi_p] \oplus F_x \cap F_y^\perp \quad \text{et} \quad F_y = [\varphi_p] \oplus F_y \cap F_x^\perp .$$

Dém: il est évident que x et y indépendants, est équivalent à

$$\Phi^2 = \sum_{k,\ell} \frac{(P(M^k \cap N^\ell) - P(M^k)P(\ell))^2}{P(M^k)\,P(N^\ell)} = 0$$

Or $\Phi^2 = Tr(P_{F_x} P_{F_y}) - 1$; d'où la proposition en utilisant 6.1.4. et l'A.C. de (F_x, F_y).

6.2.6. Définition

Si x *et* y *sont irréductibles, nous appellerons " indice de corrélation", de* (x,y) *le scalaire*

$$\alpha_{x,y} = \frac{Tr(P_{F_x} P_{F_y}) - 1}{\max(\dim F_x, \dim F_y) - 1} .$$

Remarquons que

$$0 \leqslant \alpha_{x,y} \leqslant 1$$

$\alpha_{x,y} = 0 \iff$ x et y indépendants

$\alpha_{x,y} = 1 \iff$ x et y équivalents (6.1.3.).

$\alpha_{x,y}$ indique l'importance de la liaison entre x et y, ou ce qui revient au même, le degré de proximité entre les potentiels de prévision F_x et F_y.

Nous avons ainsi, à travers ce formalisme, un point de vue original, sur le degré de liaison (ou corrélation) entre deux paramètres irréductibles.

6.3. Paramètres réductibles et indice partiel de corrélation.

6.3.1. Définition

Si u *et* v *sont les représentations associées aux paramètres qualitatifs correspondants à* x *et* y, *alors, réduire* (x,y) *signifiera réduire* u+v.

Plus précisément si

$$(M^k)_{k \in J_s} \vee (N^\ell)_{\ell \in J_t} = (L^m)_{m \in J_z}$$

alors, on voit que

$$\forall m \in J_z, \ \exists k_m, \exists \ell_m \ \text{ tels que } L^m = \bigcup_{i \in J_{k_m}} M^i = \bigcup_{j \in J_{\ell_m}} N^j$$

De plus,

$$\forall m\in J_Z\,,\ ((\varphi_{M^i})_{i\in J_{k_m}}\ ,\ (\varphi_{N^j})_{j\in J_{\ell_m}})\ \text{est irréductible}\,.$$

Notons $F^m_x = [\ \{\ \varphi_{M^i} \mid i\in J_{k_m}\ \}\ ]$

$F^m_y = [\ \{\ \varphi_{N^j} \mid j\in J_{\ell_m}\ \}\ ]$.

Il est évident que

$$F_x = \bigoplus_{m\in J_Z} F^m_x \qquad , F_y = \bigoplus_{m\in J_Z} F^m_y$$

et $$F_x + F_y = \bigoplus_{m\in J_Z} (F^m_x + F^m_y)\,.$$

On appellera " indice partiel de corrélation " le scalaire

$$\alpha^m_{x,y} = \frac{\operatorname{Tr}(P_{F^m_x} P_{F^m_y}) - 1}{\max(\dim F^m_x, \dim F^m_y) - 1}\,.$$

C'est l'indice de corrélation entre

$$(\varphi_{M^i})_{i\in J_{k_m}} \quad \text{et} \quad (\varphi_{N^j})_{j\in J_{\ell_m}}$$

en conditionnant toutes les probabilités par L^m .

(i.e. en remplaçant p_i par $q_i = p_i/P(L^m)$) .

6.3.2. Définition

On appellera " indice de corrélation " de (x,y) *le scalaire*

$$\alpha_{x,y} = \sum_{m\in J_Z} P(L^m)\ \alpha^m_{x,y}\,.$$

Remarquons que

$$0 \leqslant \alpha_{x,y} \leqslant 1$$

$$\alpha_{x,y} = 1 \iff x \text{ équivalent à } y$$

$$\alpha_{x,y} = 0 \iff \forall m \in J_z \ , \ \alpha^m_{x,m} = 0$$

$$\iff Tr(P_{F_x} P_{F_y}) = \dim(F_x \cap F_y) = z$$

$$\iff F_x = F_x \cap F_y \oplus F_x \cap F_y^{\perp}$$

$$\text{et} \qquad F_y = F_x \cap F_y \oplus F_y \cap F_x^{\perp}$$

$$\iff (\forall k \in J_s \ , \forall \ell \in J_t \ , M^k \cap N^\ell \neq \emptyset$$

$$\Rightarrow \exists \ m \text{ t.q. } M^k \cup N^\ell \subset L^m \text{ et } P(L^m)\, P(M^k \cap N^\ell) = P(M^k)\, .\, P(N^\ell))\, .$$

Nous dirons alors que x et y sont " partiellement indépendants " . Cette définition est une extension de la notion classique d'indépendance (6.2.3.), où (x,y) est irréductible . $\alpha_{x,y}$ représente comme dans 6.2.6. le degré de proximité entre F_x et F_y, $F_x \cap F_y$ étant fixé .

6.4. Analyse factorielle des correspondances (A.F.C.)

6.4.1. Définition

Soit u *et* v *les représentations associées aux paramètres qualitatifs correspondants à* x *et* y *. On appelle* A.F.C. , l'A.C. *de* (u,v) .

6.4.2. Remarque

La proximité des différentes modalités de x et y peut-être mise en évidence par l'A.C.P. de $u + v$.

$$\alpha_{x,y} = 0 \iff \forall m \in J_z \ , \ \alpha^{m}_{x,m} = 0$$

$$\iff Tr(P_{F_x} P_{F_y}) = \dim(F_x \cap F_y) = z$$

$$\iff F_x = F_x \cap F_y \oplus F_x \cap F_y^{\perp}$$

$$\text{et} \qquad F_y = F_x \cap F_y \oplus F_y \cap F_x^{\perp}$$

$$\iff (\forall k \in J_s \ , \forall \ell \in J_t \ , M^k \cap N^\ell \neq \emptyset$$

$$\Rightarrow \exists\, m \text{ t.q. } M^k \cup N^\ell \subset L^m \text{ et } P(L^m)\, P(M^k \cap N^\ell) = P(M^k)\,.\,P(N^\ell))\,.$$

Nous dirons alors que x et y sont " partiellement indépendants ". Cette définition est une extension de la notion classique d'indépendance (6.2.3.), où (x,y) est irréductible. $\alpha_{x,y}$ représente comme dans 6.2.6. le degré de proximité entre F_x et F_y, $F_x \cap F_y$ étant fixé.

6.4. Analyse factorielle des correspondances (A.F.C.)

6.4.1. Définition

Soit u *et* v *les représentations associées aux paramètres qualitatifs correspondants à* x *et* y. *On appelle* A.F.C., l'A.C. *de* (u,v).

6.4.2. Remarque

La proximité des différentes modalités de x et y peut-être mise en évidence par l'A.C.P. de u + v.