

ISSN 0248-3289

**PUBLICATIONS
DU
LABORATOIRE
DE
STATISTIQUE
ET
PROBABILITES**

N° 02-84
Décembre 1984

ETUDE COMPARATIVE DES REPRESENTATIONS GRAPHIQUES
EN ANALYSES FACTORIELLES DES CORRESPONDANCES
SIMPLES ET MULTIPLES

Alain BACCINI

**UNIVERSITE PAUL SABATIER
TOULOUSE**

ETUDE COMPARATIVE DES REPRESENTATIONS GRAPHIQUES
EN ANALYSES FACTORIELLES DES CORRESPONDANCES
SIMPLES ET MULTIPLES

Alain BACCINI

UNIVERSITE PAUL SABATIER
Laboratoire de Statistique et Probabilités
Equipe de Recherche Associée au CNRS
118, route de Narbonne
31062 TOULOUSE CEDEX

N° 02-84
Décembre 1984

S O M M A I R E

Introduction	1
1. Rappels sur l'Analyse en Composantes Principales	2
2. L'Analyse Factorielle des Correspondances (simple)	6
3. A.F.C. du tableau disjonctif complet relatif à deux variables qualitatives	16
4. A.F.C. du tableau de BURT relatif à deux variables qualitatives	27
5. A.F.C. d'un tableau disjonctif complet relatif à plus de deux variables qualitatives	33
6. A.F.C. d'un tableau de BURT relatif à plus de deux variables qualitatives	40
7. Applications de l'Analyse Canonique à l'A.F.C.	46
8. Applications de l'Analyse Canonique Généralisée (au sens de CARROLL) à l'A.F.C. multiple	54
9. Bilan	61

INTRODUCTION

Cet article constitue le développement d'une conférence donnée au Centre for Applied Statistics de l'Université de Lancaster (Royaume-Uni), en mai 1984.

Il s'inscrit dans le cadre de l'A.T.P.-C.N.R.S. "Comparaisons internationales" développée conjointement par ce Centre et le Laboratoire de Statistique et Probabilités de l'Université Paul Sabatier de Toulouse.

L'objectif de l'article est d'explicitier le principe des représentations graphiques en Analyses des Correspondances (simples et multiples) et, plus particulièrement, de faire le lien entre les représentations graphiques associées à un tableau de données qualitatives multidimensionnelles, selon que l'on utilise l'Analyse des Correspondances sur tableau disjonctif complet, la même méthode sur tableau de BURT, ou la généralisation de l'Analyse Canonique proposée par J.D. CARROLL.

L'explicitation de ces diverses représentations graphiques nous paraît fondamentale si l'on veut d'une part être en mesure de choisir la méthode la plus appropriée pour traiter un problème donné, d'autre part interpréter le plus correctement possible les sorties graphiques de chaque méthode.

Nota : Tous les graphiques ont été réalisés sur des axes orthonormés munis de la même échelle, à l'exception des deux graphiques du paragraphe 1. .

1. RAPPELS SUR L'ANALYSE EN COMPOSANTES PRINCIPALES

L'objet de ce paragraphe est de rappeler les résultats de l'Analyse en Composantes Principales (A.C.P.) nécessaires à l'Analyse Factorielle des Correspondances (A.F.C.). Ces résultats étant très classiques, nous ne les démontrons pas ; on pourra notamment se reporter à [1], [2], [3] ou [8].

1.1. Cas simple.

• Données et notations : On considère p variables quantitatives notées $Y^1, \dots, Y^j, \dots, Y^p$ observées sur n individus (ou unités statistiques) notés $1, \dots, \omega, \dots, n$, ces individus étant affectés du même poids $\frac{1}{n}$.

On pose $X^j = Y^j - \bar{y}^j$, où $\bar{y}^j = \frac{1}{n} \sum_{\omega=1}^n y_{\omega}^j$; $y_{\omega}^j = Y^j(\omega)$ est l'observation de la j -ième variable sur le ω -ième individu ; \bar{y}^j est la moyenne arithmétique empirique de Y^j ; X^j est la variable centrée associée à Y^j .

Soit X la matrice $n \times p$ telle que $(X)_{\omega}^j = x_{\omega}^j = y_{\omega}^j - \bar{y}^j$;

X est la matrice des données centrées.

Soit $V = \frac{1}{n} {}^t X X$, la matrice carrée d'ordre p des variances-covariances.

• Problème : On souhaite remplacer X par une matrice C , $n \times p$, linéairement équivalent (chaque colonne de C sera une combinaison linéaire des colonnes de X) et contenant un maximum d'"information" dans ses premières colonnes, cette information étant non redondante. La notion d'information maximum se traduit par le critère de variance maximum (il s'agit d'un cas particulier du critère des moindres carrés) ; la notion de non redondance se traduit par la contrainte de non corrélation entre les colonnes de C ; en imposant en plus une contrainte de normalisation aux combinaisons linéaires, on obtient le problème suivant :

déterminer successivement les p combinaisons linéaires C^k des variables X^j , s'écrivant $C^k = \sum_{j=1}^p u_j^k X^j$, chacune de variance maximum, sous les contraintes

$$(i) \quad \forall k = 1, \dots, p : \sum_{j=1}^p (u_j^k)^2 = 1 \text{ (normalisation) ,}$$

$$(ii) \quad \forall k = 2, \dots, p : \sum_{j=1}^p u_j^k u_j^{\ell} = 0, \quad \forall \ell = 1, \dots, k-1 \text{ (non corrélation).}$$

(k ≠ ℓ)

• Solution : Les vecteurs u^k ($u^k = t_{[u_1^k, \dots, u_p^k]}$; $k=1, \dots, p$) sont les vecteurs propres orthonormés de V , respectivement associés aux valeurs propres rangées par ordre décroissant (ces valeurs propres étant toutes positives ou nulles) ; les "variables" C^k qui leur sont associées sont appelées les *composantes principales* (ou parfois, les facteurs).

• Représentation graphique des individus : Soit C la matrice $n \times p$ donnant dans sa k -ième colonne les valeurs de C^k sur les n individus observés ; pour tout entier s ($1 \leq s \leq p-1$), soit C_s la matrice obtenue en prenant les s premières colonnes de C (il s'agit de la "meilleure approximation" de X par une matrice $n \times s$). En considérant s axes (droites orientées et munies d'une échelle) deux-à-deux perpendiculaires (appelés *axes principaux*) et en représentant chaque individu ω au moyen des valeurs respectives de la ω -ième ligne de C_s , on obtient la représentation des individus en dimension s la plus "fidèle" au tableau initial X . En A.C.P., on réalise le plus souvent ce graphique pour $s=2$ ou $s=3$.

• Expression de C : Soit U la matrice carrée d'ordre p dont la k -ième colonne est u^k ; on a : $C = XU$.

• Résultats complémentaires :

- la qualité globale du graphique réalisé en dimension s se mesure par le

rapport
$$r_s = \frac{\sum_{k=1}^s \lambda_k}{\sum_{k=1}^p \lambda_k} , \text{ où } \lambda_k \text{ est la}$$

k -ième plus grande valeur propre de V (r_s est souvent exprimé en pourcentage) ; on peut en effet vérifier : $\text{Var}(C^k) = \frac{1}{n} \sum_{\omega=1}^n (c_{\omega}^k)^2 = \lambda_k$, où $c_{\omega}^k = (C)_{\omega}^k$, et

$$\sum_{k=1}^p \lambda_k = \sum_{k=1}^p \text{Var}(C^k) = \sum_{j=1}^p \text{Var}(X^j) (= \text{tr}(V)) , \text{ trace de la matrice } V ;$$

- la qualité de la représentation de chaque individu ω dans le graphique en dimension s se mesure par le cosinus de l'angle formé dans \mathbb{R}^p par le vecteur représentant ω et sa projection sur le sous-espace vectoriel de dimension s engendré par les s premiers axes principaux ; on peut vérifier que ce cosinus vaut

$$\sqrt{\frac{\sum_{k=1}^s (c_{\omega}^k)^2}{\sum_{k=1}^p (c_{\omega}^k)^2}} ;$$

- il est aussi possible de représenter chaque variable X^j en projection dans le sous-espace vectoriel de \mathbb{R}^n de dimension s ($1 \leq s \leq p-1$) engendré par les s premières composantes principales ; il suffit pour cela de considérer les s premières colonnes de la matrice $UL^{1/2}$, où $L = \text{diag}(\lambda_1, \dots, \lambda_p)$, et d'utiliser pour chaque variable X^j les s éléments de la j -ième ligne de cette matrice.

1.2. Exemple.

L'exemple ci-dessous est adapté de [7] ; très élémentaire, il présente l'avantage de permettre de faire tous les calculs "à la main". Il correspond à 3 variables observées sur 6 individus (par exemple les notes dans 3 matières obtenues par 6 étudiants).

$$\begin{array}{l} \text{données initiales :} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \end{array} \begin{array}{ccc} \gamma^1 & \gamma^2 & \gamma^3 \\ \begin{bmatrix} 0 & 7 & 3 \\ 4 & 6 & 6 \\ 6 & 8 & 8 \\ 10 & 8 & 4 \\ 8 & 6 & 2 \\ 8 & 1 & 1 \end{bmatrix} \end{array} ; \text{ matrice } X : \begin{bmatrix} -6 & 1 & -1 \\ -2 & 0 & 2 \\ 0 & 2 & 4 \\ 4 & 2 & 0 \\ 2 & 0 & -2 \\ 2 & -5 & -3 \end{bmatrix} .$$

$$\bar{y}^1 = \bar{y}^2 = 6 \quad ; \quad \bar{y}^3 = 4 .$$

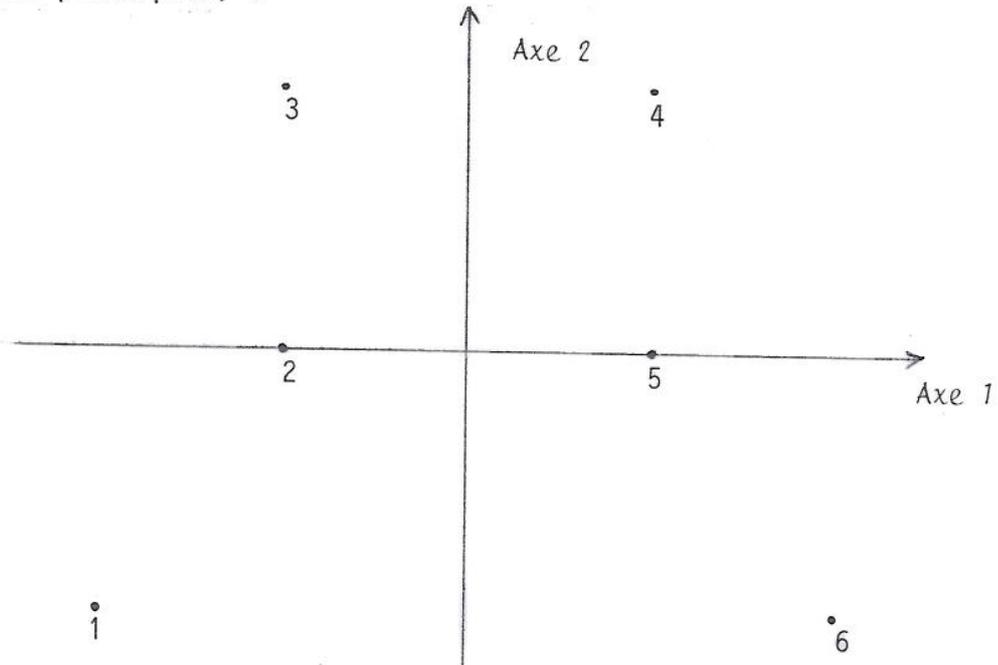
$$V = \frac{1}{6} t_{XX} = \frac{1}{6} \begin{bmatrix} 64 & -8 & -8 \\ -8 & 34 & 22 \\ -8 & 22 & 34 \end{bmatrix} .$$

$$\text{tr}(V) = 22 \quad ; \quad \lambda_1 = 12 \quad ; \quad \lambda_2 = 8 \quad ; \quad \lambda_3 = 2 .$$

$$U \approx \begin{bmatrix} 0,82 & 0,58 & 0 \\ -0,41 & 0,58 & 0,71 \\ -0,41 & 0,58 & 0,71 \end{bmatrix} .$$

$$C \approx \begin{bmatrix} -4,90 & -3,46 & 1,41 \\ -2,45 & 0 & -1,41 \\ -2,45 & 3,46 & -1,41 \\ 2,45 & 3,46 & 1,41 \\ 2,45 & 0 & 1,41 \\ 4,90 & -3,46 & -1,41 \end{bmatrix} .$$

Représentation des individus dans le *plan principal* (plan engendré par les deux premiers axes principaux) :

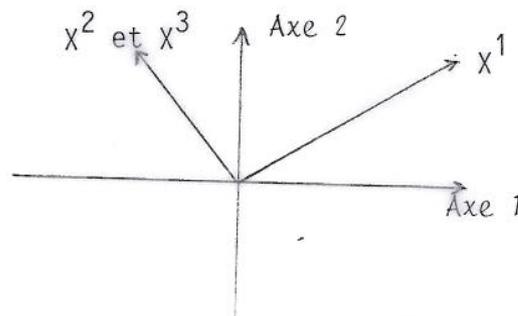


$$r_2 = \frac{20}{22} \times 100 \approx 91\%.$$

Qualité de la représentation de l'individu 1 dans le plan principal : 0,973 ($\approx 97\%$).

$$UL^{1/2} = \begin{bmatrix} 2,83 & 1,63 & 0 \\ -1,41 & 1,63 & 1 \\ -1,41 & 1,63 & -1 \end{bmatrix} .$$

Représentation des variables dans le sous-espace vectoriel de \mathbb{R}^6 engendré par C^1 et C^2 :



1.3. Cas général.

Le cas le plus général de l'Analyse en Composantes Principales d'un tableau de données (voir notamment [3] et [5]) consiste

- d'une part à attribuer un poids p_ω quelconque à chaque individu

$$(\forall \omega=1, \dots, n : p_\omega > 0 ; \sum_{\omega=1}^n p_\omega = 1),$$

• d'autre part à considérer un produit scalaire quelconque sur l'espace vectoriel \mathbb{R}^p .

Si l'on note D la matrice $\text{diag}(p_1, \dots, p_n)$ et M la matrice carrée d'ordre p , symétrique, définie-positive, définissant le produit scalaire de \mathbb{R}^p relativement à la base canonique, la matrice des variances-covariances s'écrit alors $V = {}^t X D X$, et la solution au problème posé en 1.1. est fournie par les p vecteurs propres M -orthonormés de la matrice VM respectivement associés aux valeurs propres rangées par ordre décroissant. En notant toujours U la matrice carrée d'ordre p de ces vecteurs propres disposés en colonnes, il vient dans ce cas :

$$C = X M U .$$

2. L'ANALYSE FACTORIELLE DES CORRESPONDANCES (SIMPLE)

Comme dans le paragraphe 1., les résultats les plus classiques de l'A.F.C. ne seront pas démontrés ici ; nous renvoyons pour cela à [2], [3] ou [8].

2.1. Données et notations.

On considère 2 variables qualitatives, X^1 à ℓ modalités et X^2 à c modalités, observées sur n individus affectés du même poids ($\frac{1}{n}$), ainsi que la table de contingence associée T , $\ell \times c$ (X^1 est disposée en lignes et X^2 en colonnes).

On note :

- n_{ij} , $1 \leq i \leq \ell$, $1 \leq j \leq c$, le terme général de T ;
- $n_{i.}$, $1 \leq i \leq \ell$, l'effectif marginal de la ligne i ; $f_{i.} = \frac{n_{i.}}{n}$;
- $n_{.j}$, $1 \leq j \leq c$, l'effectif marginal de la colonne j ; $f_{.j} = \frac{n_{.j}}{n}$;
- $D_1 = \text{diag}(f_{1.}, \dots, f_{\ell.})$;
- $D_2 = \text{diag}(f_{.1}, \dots, f_{.c})$;
- A_1 la matrice $\ell \times c$ des profils-lignes disposés en lignes ;

on a : $A_1 = \frac{1}{n} D_1^{-1} T$;

- A_2 la matrice $c \times \ell$ des profils-colonnes disposés en lignes ;

on a : $A_2 = \frac{1}{n} D_2^{-1} {}^t T$;

2.2. Principe de l'A.F.C. vue comme une double A.C.P.

On veut étudier les liaisons entre modalités de X^1 et modalités de X^2 ; pour cela, on réalise une A.C.P. des modalités de X^1 représentées par les profils-lignes pondérés par les $f_{i.}$, puis une A.C.P. des modalités de X^2 représentées par les

profils-colonnes pondérés par les $f_{.j}$; on compare ensuite les deux ensembles de résultats (se reporter à [2] ou [8]). Dans chaque cas, la métrique M introduite en 1.3. est la métrique du khi-deux.

2.3. A.C.P. des profils-lignes.

• Matrice des données : A_1 ; les données ne seront pas centrées car on peut vérifier qu'il est équivalent de centrer ou de ne pas le faire, et qu'il est plus simple ici de ne pas centrer (voir la remarque plus loin).

• Matrice des poids : D_1 .

• Métrique sur \mathbb{R}^c : C'est la métrique du χ^2 , associée à la matrice D_2^{-1} sur la base canonique.

• Matrice à diagonaliser : C'est VM en A.C.P., où $V = {}^t X D X$;

ici : X , $n \times p$, est remplacée par A_1 , $l \times c$;

D , $n \times n$, est remplacée par D_1 , $l \times l$;

M , $p \times p$, est remplacée par D_2^{-1} , $c \times c$;

donc, VM est remplacée par ${}^t A_1 D_1 A_1 D_2^{-1}$;

or, $D_1 A_1 D_2^{-1} = \frac{1}{n} D_1 D_1^{-1} T D_2^{-1} = \frac{1}{n} T D_2^{-1} = {}^t A_2$; on en déduit que la matrice à diagonaliser est ${}^t A_1 {}^t A_2$.

• Propriétés : Appelons $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{c-1}$ les valeurs propres de ${}^t A_1 {}^t A_2$ rangées par ordre décroissant.

→ $\forall k=0,1,\dots,c-1 : 0 \leq \lambda_k \leq 1$;

→ $\lambda_0 = 1$;

→ le vecteur propre D_2^{-1} -normé associé

à λ_0 s'écrit :

$$[u_2^0] = \begin{bmatrix} f_{.1} \\ \vdots \\ f_{.j} \\ \vdots \\ f_{.c} \end{bmatrix} ;$$

Remarque : La j -ième variable de cette A.C.P. (j -ième colonne de A_1) a pour moyenne $f_{.j}$; u_2^0 est donc le vecteur des moyennes de ces variables ; on peut vérifier qu'aux vecteurs propres de ${}^tA_1 {}^tA_2 D_2^{-1}$ -orthogonaux à u_2^0 sont associées des composantes principales centrées ; au lieu de centrer les données, on se contente donc de retenir les vecteurs propres à partir du rang 1 (D_2^{-1} -orthogonaux à u_2^0), les résultats étant identiques.

→ si $\ell \leq c$ (ce que l'on peut supposer), on a :

$$\lambda_\ell = \lambda_{\ell+1} = \dots = \lambda_{c-1} = 0.$$

• Matrices des vecteurs propres et des valeurs propres :

Pour les raisons indiquées ci-dessus, on ne retient que les vecteurs propres de rang compris entre 1 et $\ell-1$; ces vecteurs propres seront indicés par 2 (en bas) ;

soit u_2^k le k -ième vecteur propre D_2^{-1} -normé de ${}^tA_1 {}^tA_2$ ($k=1, \dots, \ell-1$) ;

on notera U_2 la matrice $c \times (\ell-1)$ de ces vecteurs (deux à deux D_2^{-1} -orthogonaux) disposés en colonnes et

$L = \text{diag}(\lambda_1, \dots, \lambda_{\ell-1})$ la matrice diagonale des valeurs propres rangées par ordre décroissant.

• Matrice des composantes principales : elle s'écrit $C = XMU$ en A.C.P. ; on obtient donc ici :

$$C_1 = A_1 D_2^{-1} U_2 ;$$

c'est cette matrice, $\ell \times (\ell-1)$, qui permettra de faire une représentation graphique des ℓ modalités de X^1 sur $\ell-1$ axes principaux.

2.4. A.C.P. des profils-colonnes.

• Matrice des données : A_2 .

• Matrice des poids : D_2 .

• Métrique sur \mathbb{R}^p : métrique du χ^2 , associée à D_1^{-1} sur la base canonique.

• Matrice à diagonaliser : ${}^tA_2 {}^tA_1$;

• Propriétés : Appelons $u_1^0, u_1^1, \dots, u_1^{\ell-1}$ les ℓ vecteurs propres D_1^{-1} -orthonormés de ${}^tA_2 {}^tA_1$ associés aux valeurs propres $\mu_0, \mu_1, \dots, \mu_{\ell-1}$ rangées par ordre décroissant ($\mu_0 \geq \mu_1 \geq \dots \geq \mu_{\ell-1}$) ; les propriétés des vecteurs propres et valeurs propres de ${}^tA_2 {}^tA_1$ sont les mêmes que celles des vecteurs propres et valeurs propres de ${}^tA_1 {}^tA_2$, et l'on a de plus :

→ $\forall k=1, \dots, \ell-1 : \mu_k = \lambda_k$ (on a aussi bien sûr : $\mu_0 = \lambda_0 = 1$) ;

→ $\forall k=1, \dots, \ell-1 : [u_1^k] = \frac{1}{\sqrt{\lambda_k}} {}^t A_2 [u_2^k]$ (à condition que λ_k soit non nulle).

• Matrices des vecteurs propres et des valeurs propres :

La matrice diagonale des valeurs propres est la même que dans le cas précédent, soit L ; notons U_1 la matrice $\ell \times (\ell-1)$ des vecteurs propres D_1^{-1} -orthonormés de ${}^t A_2 {}^t A_1$ disposés en colonnes ; on a donc : $U_1 = {}^t A_2 U_2 L^{-1/2}$.

• Matrice des composantes principales :

Soit C_2 cette matrice $c \times (\ell-1)$; on a : $C_2 = A_2 D_1^{-1} U_1$.

La représentation graphique des c modalités de X^2 sur les $\ell-1$ axes principaux se fait au moyen des lignes de C_2 .

2.5. Relations entre les deux A.C.P.

• Formules de transition :

$$\begin{aligned} \text{On a : } C_2 &= A_2 D_1^{-1} U_1 = A_2 D_1^{-1} {}^t A_2 U_2 L^{-1/2} = \frac{1}{n} A_2 D_1^{-1} T D_2^{-1} U_2 L^{-1/2} \\ &= A_2 A_1 D_2^{-1} U_2 L^{-1/2} = A_2 C_1 L^{-1/2} . \end{aligned}$$

$$\left. \begin{aligned} \text{Donc : } C_2 &= A_2 C_1 L^{-1/2} \\ \text{et de même : } C_1 &= A_1 C_2 L^{-1/2} \end{aligned} \right\} \text{ formules de transition.}$$

$$\text{Elles s'écrivent encore : } (C_2)_j^k = \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{\ell} (A_2)_j^i (C_1)_i^k = \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{\ell} \frac{n_{ij}}{n_{.j}} (C_j)_i^k ;$$

$$(C_1)_i^k = \frac{1}{\sqrt{\lambda_k}} \sum_{j=1}^c \frac{n_{ij}}{n_{.i}} (C_2)_j^k .$$

• Conséquence : représentations simultanées brutes et barycentriques :

L'homogénéité des matrices C_1 et C_2 (matrices de composantes principales de deux A.C.P. de profils issus du même tableau T), ainsi que les formules de transition entre ces deux matrices, autorisent à confondre les axes principaux des deux A.C.P. et à représenter simultanément les modalités de X^1 et celles de X^2 ;

En posant $\Gamma_1 = C_1 L^{1/2} = A_1 C_2$ et $\Gamma_2 = C_2 L^{1/2} = A_2 C_1$, on peut alors faire trois représentations graphiques :

→ si l'on représente les modalités X_i^1 de X^1 au moyen des lignes de C_1 et les modalités X_j^2 de X^2 au moyen des lignes de C_2 , on obtient la *représentation simultanée brute*, dans laquelle on n'a aucune relation barycentrique ;

→ si l'on représente les X_i^1 au moyen des lignes de Γ_1 et les X_j^2 au moyen de celles de C_2 , chaque X_i^1 est barycentre de l'ensemble des X_j^2 , chaque X_j^2 étant affecté du poids $\frac{n_{ij}}{n_i}$ (l'ensemble de ces pondérations constituant le i -ième profil-ligne associé à X_i^1) ; c'est la *première représentation barycentrique* ;

→ si l'on représente les X_i^1 au moyen des lignes de C_1 et les X_j^2 au moyen de celles de Γ_2 , chaque X_j^2 est barycentre de l'ensemble des X_i^1 , chaque X_i^1 étant affecté du poids $\frac{n_{ij}}{n_j}$; c'est la *seconde représentation barycentrique*.

Remarque : les logiciels d'A.F.C. ne réalisent en général que la représentation brute ; il est fréquent que l'on rajoute ensuite "à la main" l'une des deux représentations barycentriques.

• Remarque importante :

u_2^k est vecteur propre de ${}^t A_1 {}^t A_2$ associé à λ_k ; en posant $a_2^k = D_2^{-1} u_2^k$, on vérifie simplement que a_2^k est vecteur propre de $A_2 A_1$ associé à λ_k ; matriciellement, on obtient donc : $A_2 A_1 D_2^{-1} U_2 = D_2^{-1} U_2 L$. En portant dans $C_2 = A_2 A_1 D_2^{-1} U_2 L^{-1/2}$ (voir ci-dessus les formules de transition), on obtient :

$$C_2 = D_2^{-1} U_2 L^{1/2} ; \text{ de la même manière, il vient :}$$

$$C_1 = D_1^{-1} U_1 L^{1/2} .$$

Ces formules seront utilisées en A.F.C. multiple.

• Formule de reconstitution des données :

On peut vérifier l'égalité :

$$f_{ij} = f_i \cdot f_j \left[1 + \sum_{k=1}^{l-1} \frac{1}{\sqrt{\lambda_k}} (C_1)_i^k (C_2)_j^k \right] .$$

Elle permet donc d'écrire f_{ij} en fonction du produit $f_i \cdot f_j$ et des matrices C_1 , C_2 et L .

2.6. Quelques résultats complémentaires.

- Propriété des valeurs propres : On peut vérifier les égalités suivantes :

$$\sum_{k=1}^{\ell-1} \lambda_k = \text{tr}({}^t A_1 {}^t A_2) - 1 = \text{tr}({}^t A_2 {}^t A_1) - 1 = \Phi^2 = \frac{\chi^2}{n} ,$$

où χ^2 est le khi-deux de contingence calculé sur la table T.

- Qualité globale des représentations graphiques : Comme en A.C.P., on mesure la qualité des représentations graphiques (brute ou barycentriques) en dimension s ($1 \leq s \leq \ell-1$) par le rapport :

$$r_s = \frac{\sum_{k=1}^s \lambda_k}{\sum_{k=1}^{\ell-1} \lambda_k} .$$

- Qualité de la représentation d'une modalité de χ^1 ou de χ^2 : Pour une représentation graphique en dimension s , on utilise là encore les résultats de l'A.C.P. en mesurant la qualité de la représentation de la i -ième modalité de χ^1 par la quantité :

$$\sqrt{\frac{\sum_{k=1}^s [(C_1)_i^k]^2}{\sum_{k=1}^{\ell-1} [(C_1)_i^k]^2}} \quad (\text{il s'agit d'un cosinus}) ;$$

on utilise l'expression correspondante pour la j -ième modalité de χ^2 .

- Contribution relative de chaque modalité à l'inertie de son nuage :

On appelle inertie du nuage des profils-lignes la quantité

$$I_1 = \sum_{i=1}^{\ell} f_i \cdot [d(x_i^1, u_2^0)]^2 , \text{ où } x_i^1 \text{ est le vecteur de } \mathbb{R}^c$$

(muni de la base canonique) associé à X_i^1 (sa j -ième coordonnée est $\frac{n_{ij}}{n_i}$), où u_2^0 (introduit en 2.3) est le barycentre des x_i^1 ($i = 1, \dots, \ell$), et où d désigne la distance au sens de la métrique du χ^2 . On peut alors vérifier :

$$[d(x_i^1, u_2^0)]^2 = \sum_{k=1}^{\ell-1} [(C_1)_i^k]^2 \quad \text{et} \quad I_1 = \Phi^2 ;$$

on en déduit que la part d'inertie due à X_i^1 s'écrit : $\frac{f_i}{\Phi^2} \sum_{k=1}^{\ell-1} [(C_1)_i^k]^2$;

on a également la quantité correspondante pour la part d'inertie due à X_j^2 dans I_2 .

Ces quantités servent à apprécier l'importance de chaque modalité dans l'A.F.C..

• Contribution relative de chaque modalité à la détermination de chaque axe principal : On appelle inertie du nuage des profils-lignes en projection sur le k-ième axe principal la quantité

$$I_1^k = \sum_{i=1}^{\ell} f_i \cdot [d(x_i^{1k}, u_2^0)]^2 \quad , \quad \text{où } x_i^{1k} \text{ est la projection}$$

D_2^{-1} -orthogonale de x_i^1 sur le k-ième axe principal (dans la mesure où l'A.F.C. réalisée est équivalente à une A.C.P. centrée, u_2^0 est l'origine de \mathbb{R}^C et est donc confondu avec sa projection).

On peut alors vérifier :

$$[d(x_i^{1k}, u_2^0)]^2 = [(C_1)_i^k]^2 \quad \text{et} \quad I_1^k = \lambda_k \quad ;$$

on en déduit que la part relative d'inertie due à X_i^1 , et donc sa contribution à la détermination de l'axe correspondant, s'écrit :

$$\frac{f_i}{\lambda_k} \cdot [(C_1)_i^k]^2 \quad ; \quad \text{on obtient la quantité correspondante pour}$$

X_j^2 et I_2^k .

Ces quantités sont utilisées pour interpréter les axes principaux de l'A.F.C..

2.7. Exemple.

Nous présentons ici aussi un exemple très simple qui sera utilisé tout au long de cet article. La table de contingence est la suivante :

	X_1^2	X_2^2	X_3^2	X_4^2
X_1^1	1	1	0	0
X_2^1	0	1	1	0
X_3^1	0	0	1	1

(T) .

$$\ell = 3 \quad , \quad c = 4 \quad , \quad n = 6.$$

Les $n_{i.}$ valent tous 2 ($i=1,2,3$) ; les $n_{.j}$ valent 1, 2, 2 et 1.

$$D_1 = \text{diag}\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3} I_3 \quad ; \quad D_1^{-1} = 3 I_3 \quad .$$

$$D_2 = \text{diag}\left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right) \quad ; \quad D_2^{-1} = \text{diag}(6, 3, 3, 6) \quad .$$

$$A_1 = \frac{1}{6} D_1^{-1} T = \frac{1}{2} T \quad ;$$

$$A_2 = \frac{1}{6} D_2^{-1} t_T = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix} \quad .$$

Matrice à diagonaliser dans l'A.C.P. des profils-lignes :

$${}^t_{A_1} \quad {}^t_{A_2} = \frac{1}{4} \begin{bmatrix} 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad .$$

$$\lambda_0 = 1 \quad \lambda_1 = 0,75 \quad \lambda_2 = 0,25 \quad \lambda_3 = 0 \quad .$$

Matrice à diagonaliser dans l'A.C.P. des profils-colonnes :

$${}^t_{A_2} \quad {}^t_{A_1} = \frac{1}{4} \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \quad .$$

$$\lambda_0 = 1 \quad \lambda_1 = 0,75 \quad \lambda_2 = 0,25 \quad .$$

$$\Phi^2 = 0,75 + 0,25 = 1 \quad ; \quad \chi^2 = 6 \quad .$$

$$U_1 \approx \begin{bmatrix} 0,41 & 0,24 \\ 0 & -0,47 \\ -0,41 & 0,24 \end{bmatrix} \quad .$$

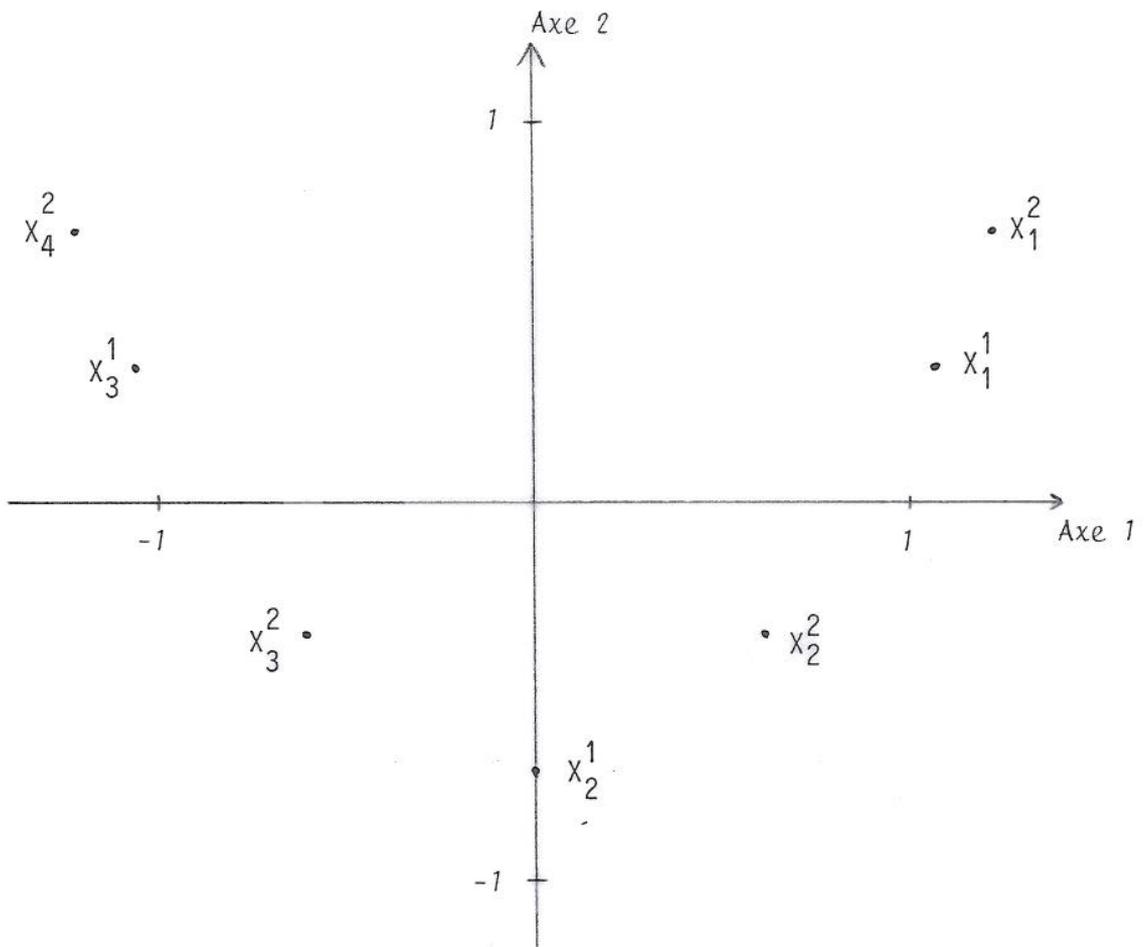
$$C_1 = D_1^{-1} U_1 L^{1/2} = \begin{bmatrix} 1,06 & 0,35 \\ 0 & -0,71 \\ -1,06 & 0,35 \end{bmatrix} \quad ;$$

$$\left. \begin{array}{l} C_2 = D_2^{-1} U_2 L^{1/2} \\ U_2 = {}^t_{A_1} U_2 L^{-1/2} \end{array} \right\} \Rightarrow C_2 = D_2^{-1} {}^t_{A_1} U_1 = \begin{bmatrix} 1,22 & 0,71 \\ 0,61 & -0,35 \\ -0,61 & -0,35 \\ -1,22 & 0,71 \end{bmatrix} \quad .$$

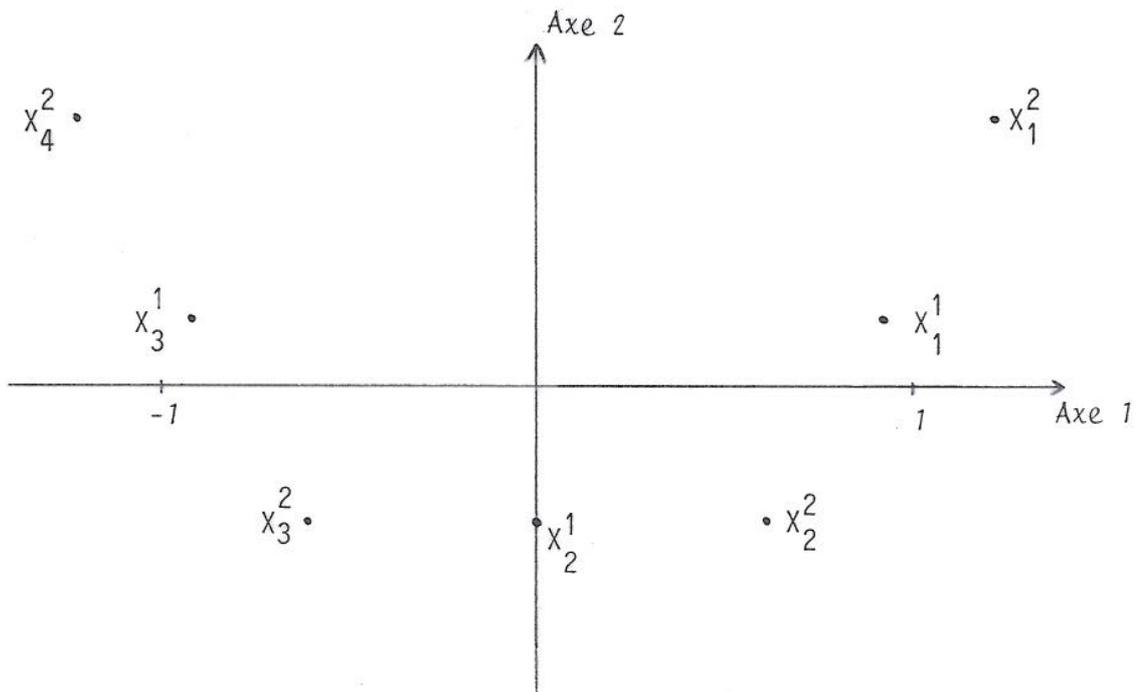
$$\Gamma_1 = C_1 L^{1/2} = \begin{bmatrix} 0,92 & 0,18 \\ 0 & -0,35 \\ -0,92 & 0,18 \end{bmatrix} \quad ; \quad \Gamma_2 = C_2 L^{1/2} = \begin{bmatrix} 1,06 & 0,35 \\ 0,53 & -0,18 \\ -0,53 & -0,18 \\ -1,06 & 0,35 \end{bmatrix} .$$

Remarque : On constate que, dans la pratique, on ne diagonalise que la matrice ${}^t A_2 {}^t A_1$.

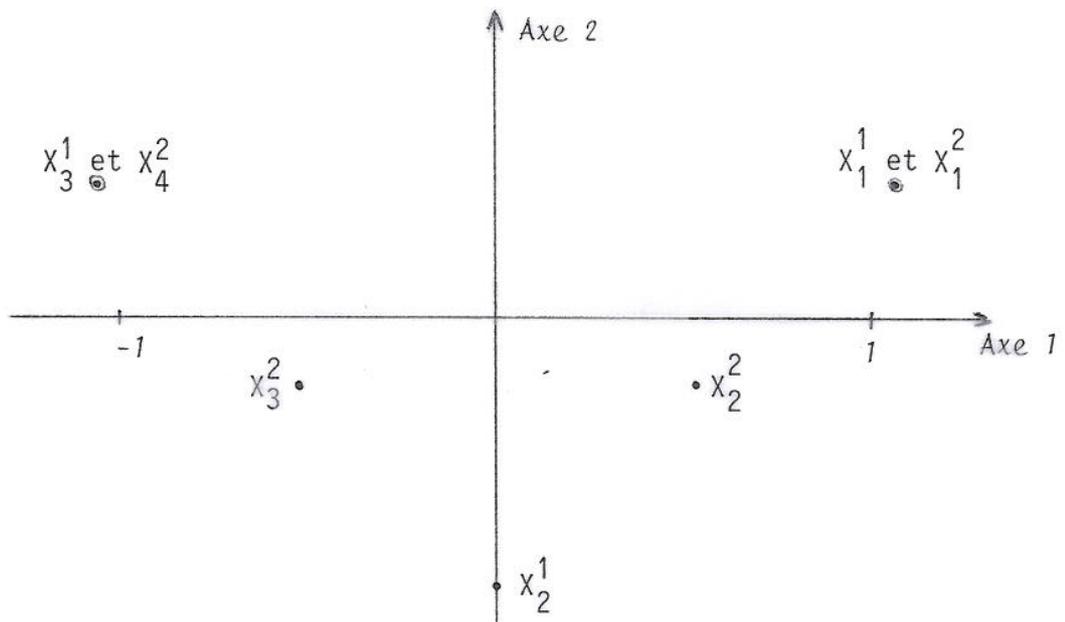
Représentation simultanée brute :



Première représentation barycentrique :



Seconde représentation barycentrique :



Toutes les mesures de qualité des graphiques en dimension 2 sont ici égales à 1 car $\ell-1 = 2$.

Les contributions relatives des modalités à l'inertie de leur nuage sont :

$$\begin{array}{l} \text{pour } X^1, \quad X_1^1 : 42\% \quad X_2^1 : 17\% \quad X_3^1 : 42\% \quad ; \\ \text{pour } X^2, \quad X_1^2 : 33\% \quad X_2^2 : 17\% \quad X_3^2 : 17\% \quad X_4^2 : 33\% \quad . \end{array}$$

Les contributions relatives des modalités à la détermination des axes sont :
pour l'axe 1 et pour X^1 ,

$$X_1^1 : 50\% \quad X_2^1 : 0\% \quad X_3^1 : 50\% \quad ;$$

pour l'axe 1 et pour X^2 ,

$$X_1^2 : 33\% \quad X_2^2 : 17\% \quad X_3^2 : 17\% \quad X_4^2 : 33\% \quad ;$$

pour l'axe 2 et pour X^1 ,

$$X_1^1 : 17\% \quad X_2^1 : 67\% \quad X_3^1 : 17\% \quad ;$$

pour l'axe 2 et pour X^2 ,

$$X_1^2 : 33\% \quad X_2^2 : 17\% \quad X_3^2 : 17\% \quad X_4^2 : 33\% \quad .$$

3. A.F.C. DU TABLEAU DISJONCTIF COMPLET RELATIF A DEUX VARIABLES QUALITATIVES

Dans ce paragraphe, comme dans le suivant, on s'intéresse à l'A.F.C. de tableaux particuliers dans le cas où l'on ne dispose que de deux variables qualitatives ; c'est la généralisation de l'A.F.C. de ces tableaux au cas de plus de deux variables qualitatives qui permet de définir l'Analyse Factorielle des Correspondances Multiple (voir les paragraphes 5. et 6.).

Nous sommes amenés à préciser ici un certain nombre d'éléments développés par ailleurs dans [6].

3.1. Définitions préliminaires.

On considère comme en 2. deux variables qualitatives X^1 , à ℓ modalités, et X^2 , à c modalités, observées simultanément sur n individus, affectés du même poids $\frac{1}{n}$.

• Variables indicatrices des modalités d'une variable qualitative :

Pour X^1 , on appelle ainsi les ℓ variables $X_{(1)}^1, \dots, X_{(i)}^1, \dots, X_{(\ell)}^1$ définies par :

$$X_{(i)}^1(\omega) = \begin{cases} 1 & , \text{ si } X^1(\omega) = X_i^1 \text{ (l'individu } \omega \text{ a présenté} \\ & \text{ la modalité } i \text{ de } X^1) ; \\ 0 & , \text{ sinon .} \end{cases}$$

Les "indicatrices" sont définies de la même manière pour toute variable qualitative, et ne prennent donc que les valeurs 1 ($n_{i.}$ fois pour $X_{(i)}^1$, si $n_{i.}$ est l'effectif marginal de la i -ième modalité de X^1) et 0 ($n-n_{i.}$ fois).

• Tableau des indicatrices :

On note X_1 le tableau $n \times \ell$ dans lequel on dispose en lignes les n individus observés et en colonnes les ℓ variables indicatrices associées à la variable X^1 ;

		$X_{(1)}^1$...	$X_{(i)}^1$...	$X_{(\ell)}^1$	
$X_1 :$	1	0	...	0	...	1	1
	⋮	1	...	0	...	0	1
	ω	0	...	1	...	0	1
	⋮	0	...	1	...	0	1
	n					n	
		$n_{1.}$...	$n_{i.}$...	$n_{\ell.}$	n

Les effectifs marginaux des lignes de ce tableau valent tous 1 ; ceux des colonnes valent $n_{1.}, \dots, n_{i.}, \dots, n_{\ell.}$; ce sont les effectifs marginaux des lignes de la table de contingence T étudiée en 2. ; ce tableau est appelé tableau des indicatrices de X^1 ; on définit pareillement le tableau $X_2, n \times c$, des indicatrices de X^2 .

• Tableau disjonctif complet :

Le tableau disjonctif complet relatif à X^1, X^2 et aux n individus observés est le tableau $X, n \times (\ell+c)$, dans lequel on dispose en lignes les individus et en colonnes les variables indicatrices des modalités de X^1 , suivies de celles de X^2 :

		$x_{(1)}^1 \cdots x_{(\ell)}^1$	$x_{(1)}^2 \cdots x_{(c)}^2$	
X :	1 ⋮ ω ⋮ n	0 ... 1	1 ... 0	2
	
		1 ... 0	0 ... 1	2
		$n_{1.} \cdots n_{\ell.}$	$n_{.1} \cdots n_{.c}$	2n

On notera que X ne comporte que des 0 et des 1, que ses effectifs marginaux des lignes valent tous 2 et qu'on retrouve comme effectifs marginaux des colonnes les $n_{i.}$ puis les $n_{.j}$; on a : $X = [X_1 \mid X_2]$.

3.2. Principe et notations.

Nous allons réaliser ici l'A.F.C. du tableau X, comme s'il s'agissait d'une table de contingence ordinaire, et essayer de préciser les résultats que l'on obtient.

Les notations introduites en 2. seront réutilisées avec la même signification ; de plus, pour désigner les nouvelles matrices définies à partir de X, nous utiliserons les mêmes lettres qu'en 2, mais en les surlignant ; nous avons ainsi :

$$\bar{T} = X \quad ; \quad \bar{D}_1 = \frac{1}{n} I_n \quad ; \quad \bar{D}_1^{-1} = n I_n \quad ;$$

$$\bar{D}_2 = \frac{1}{2n} \text{diag} (n_{1.}, \dots, n_{\ell.}, n_{.1}, \dots, n_{.c}) = \frac{1}{2} \left[\begin{array}{c|c} \bar{D}_1 & 0 \\ \hline 0 & \bar{D}_2 \end{array} \right] \quad ;$$

$$\bar{D}_2^{-1} = 2 \left[\begin{array}{c|c} \bar{D}_1^{-1} & 0 \\ \hline 0 & \bar{D}_2^{-1} \end{array} \right] \quad ;$$

$$\bar{A}_1 = \frac{1}{2n} \bar{D}_1^{-1} \bar{T} = \frac{1}{2} X \quad ;$$

$$\bar{A}_2 = \frac{1}{2n} \bar{D}_2^{-1} t_{\bar{T}} = \frac{1}{n} \left[\begin{array}{c|c} \bar{D}_1^{-1} & 0 \\ \hline 0 & \bar{D}_2^{-1} \end{array} \right] \begin{bmatrix} t_{X_1} \\ t_{X_2} \end{bmatrix} = \frac{1}{n} \left[\begin{array}{c|c} \bar{D}_1^{-1} t_{X_1} \\ \hline \bar{D}_2^{-1} t_{X_2} \end{array} \right] .$$

3.3. A.C.P. des profils-lignes de X.

• Matrice à diagonaliser :

$$\text{C'est } t_{\bar{A}_1} \quad t_{\bar{A}_2} = \frac{1}{2n} \left[\begin{array}{c} t_{X_1} \\ t_{X_2} \end{array} \right] [X_1 \bar{D}_1^{-1} \mid X_2 \bar{D}_2^{-1}]$$

$$= \frac{1}{2n} \left[\begin{array}{c|c} t_{X_1 X_1} D_1^{-1} & t_{X_1 X_2} D_2^{-1} \\ \hline t_{X_2 X_1} D_1^{-1} & t_{X_2 X_2} D_2^{-1} \end{array} \right] = \frac{1}{2n} \left[\begin{array}{c|c} n I_\ell & T D_2^{-1} \\ \hline t_T D_1^{-1} & n I_c \end{array} \right] = \frac{1}{2} \left[\begin{array}{c|c} I_\ell & t_{A_2} \\ \hline t_{A_1} & I_c \end{array} \right] .$$

• Propriétés des éléments propres de $t_{\bar{A}_1} t_{\bar{A}_2}$:

Soit v l'un quelconque des vecteurs propres \bar{D}_2^{-1} -normé de cette matrice et soit ν la valeur propre associée ; v est un vecteur de $\mathbb{R}^{\ell+c}$ que l'on peut écrire

$$[v] = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} , \text{ où } v_1 \in \mathbb{R}^\ell \text{ et } v_2 \in \mathbb{R}^c .$$

$t_{\bar{A}_1} t_{\bar{A}_2} v = \nu v$ peut encore s'écrire :

$$\begin{cases} v_1 + t_{A_2} v_2 = 2\nu v_1 & \implies t_{A_2} v_2 = (2\nu-1)v_1 & (1) \quad ; \\ t_{A_1} v_1 + v_2 = 2\nu v_2 & \implies t_{A_1} v_1 = (2\nu-1)v_2 & (2) \quad ; \end{cases}$$

prémultipliant (1) par t_{A_1} , il vient :

$$t_{A_1} t_{A_2} v_2 = (2\nu-1) t_{A_1} v_1 = (2\nu-1)^2 v_2 \text{ (d'après (2))} ;$$

on obtient de même : $t_{A_2} t_{A_1} v_1 = (2\nu-1)^2 v_1$.

v_1 est donc un vecteur propre de $t_{A_2} t_{A_1}$, v_2 un vecteur propre de $t_{A_1} t_{A_2}$ et $(2\nu-1)^2$ est une valeur propre commune à ces deux matrices.

Comme réciproquement on peut montrer que l'on génère l'ensemble des vecteurs propres de $t_{\bar{A}_1} t_{\bar{A}_2}$ à partir de ceux de $t_{A_1} t_{A_2}$ et de ceux de $t_{A_2} t_{A_1}$, on a les résultats suivants :

Soient $\lambda_0 = 1 \geq \lambda_1 \geq \dots \geq \lambda_{\ell-1}$ les valeurs propres non nulles de

$t_{A_1} t_{A_2}$ et de $t_{A_2} t_{A_1}$ ($t_{A_1} t_{A_2}$ a en plus $c-\ell$ valeurs propres nulles).

Les $\ell+c$ valeurs propres de $t_{\bar{A}_1} t_{\bar{A}_2}$, ν_k , sont définies par :

$$\nu_k = \frac{1 \pm \sqrt{\lambda_k}}{2} ;$$

→ à $\lambda_0 = 1$ est d'abord associé $\nu_0 = 1$, un vecteur propre étant $\begin{bmatrix} u_0^1 \\ 0 \\ u_0^2 \end{bmatrix}$,

si $[u_1^0] = \begin{bmatrix} \vdots \\ f \\ \vdots \\ i \\ \vdots \end{bmatrix}$ et $[u_2^0] = \begin{bmatrix} \vdots \\ f \\ \vdots \\ j \\ \vdots \end{bmatrix}$; ce vecteur propre est sans

intérêt dans cette analyse et ne sera pas pris en compte par la méthode ;

→ aux valeurs propres λ_k vérifiant $1 > \lambda_k > 0$ ($k=1, \dots, \ell-1$) sont associées dans un premier temps les valeurs propres $v_k = \frac{1+\sqrt{\lambda_k}}{2}$; on définit ainsi $\ell-1$ valeurs propres rangées par ordre décroissant et comprises entre 1 et $\frac{1}{2}$; les vecteurs propres associés sont de la forme $\begin{bmatrix} u_1^k \\ u_2^k \end{bmatrix}$ où u_1^k et u_2^k sont les vecteurs propres respectifs de ${}^t A_2$ ${}^t A_1$ et de ${}^t A_1$ ${}^t A_2$ associés à λ_k .

→ à l'ensemble des $c-\ell$ valeurs propres nulles de ${}^t A_1$ ${}^t A_2$ est associé un ensemble de $c-\ell$ valeurs propres de ${}^t \bar{A}_1$ ${}^t \bar{A}_2$ égales à $\frac{1}{2}$, notées v_ℓ, \dots, v_{c-1} ; les vecteurs propres sont de la forme $\begin{bmatrix} 0 \\ u_k \end{bmatrix}$ si ${}^t A_1$ ${}^t A_2$ $u_k = 0$;

→ aux λ_k vérifiant $1 > \lambda_k > 0$ sont ensuite associées les $\ell-1$ valeurs propres $v_k = \frac{1-\sqrt{\lambda_k}}{2}$ ($k=c, \dots, c+\ell-2$) comprises entre $\frac{1}{2}$ et 0 ; les vecteurs propres sont de la forme $\begin{bmatrix} u_1^{c+\ell-1-k} \\ -u_2^{c+\ell-1-k} \end{bmatrix}$.

→ à $\lambda_0 = 1$ est enfin associée la valeur propre $v_{c+\ell-1} = 0$ correspondant au vecteur propre $\begin{bmatrix} u_1^0 \\ -u_2^0 \end{bmatrix}$.

Seuls les vecteurs propres du rang 1 au rang $\ell-1$ ont un intérêt pour l'analyse ; on notera \bar{U}_2 la matrice $(\ell+c) \times (\ell-1)$ de ces vecteurs \bar{D}_2^{-1} -orthonormés disposés en colonnes.

Remarque : On peut vérifier : $\sum_{k=1}^{\ell+c-1} v_k = \frac{\ell+c-2}{2}$.

Le rapport $\bar{r}_k = \frac{v_k}{\sum_{k=1}^{\ell+c-1} v_k} = \frac{1+\sqrt{\lambda_k}}{\ell+c-2}$ n'a pas de signification concrète comme

c'était le cas en A.F.C.. Toutefois on pourrait dans ce cas calculer

$$\frac{[(\ell+c-2)\bar{r}_k - 1]^2}{\phi^2}, \text{ en déterminant au préalable } \phi^2.$$

• Matrice des vecteurs propres :

D'après ce qui précède, on sait que tout vecteur propre \bar{D}_2^{-1} -normé de \bar{U}_2 s'écrit

$[u] = \alpha \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, où u_1 est un vecteur propre D_1^{-1} -normé de ${}^t A_2$ ${}^t A_1$, u_2 le vecteur propre D_2^{-1} -normé de ${}^t A_1$ ${}^t A_2$ correspondant et α un certain réel ; il vient :

$$\begin{aligned} \|u\|_{D_2^{-1}}^2 (=1) &= {}^t[u] \bar{D}_2^{-1} [u] = 2\alpha^2 \left[{}^t[u_1] D_1^{-1} \mid {}^t[u_2] D_2^{-1} \right] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= 2\alpha^2 \left[{}^t[u_1] D_1^{-1} [u_1] + {}^t[u_2] D_2^{-1} [u_2] \right] = 4\alpha^2 ; \end{aligned}$$

on a donc $\alpha = \pm \frac{1}{2}$, et nous prendrons $\alpha = + \frac{1}{2}$.

On peut donc maintenant écrire : $\bar{U}_2 = \frac{1}{2} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$.

• Matrice des composantes principales :

Soit \bar{C}_1 cette matrice ; elle s'écrit : $\bar{C}_1 = \bar{A}_1 \bar{D}_2^{-1} \bar{U}_2$,

$$\begin{aligned} \text{soit } \bar{C}_1 &= \frac{1}{2} \times \left[\begin{array}{c|c} D_1^{-1} & 0 \\ \hline 0 & D_2^{-1} \end{array} \right] \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \\ &= \frac{1}{2} [X_1 D_1^{-1} U_1 + X_2 D_2^{-1} U_2] . \end{aligned}$$

Nous commenterons l'utilisation de cette matrice au point 3.4..

3.4. A.C.P. des profils-colonnes de X.

• Matrice à diagonaliser et matrice des vecteurs propres :

$$\begin{aligned} \text{Il s'agit de } {}^t \bar{A}_2 \bar{A}_1 &= \frac{1}{2n} \left[X_1 D_1^{-1} \mid X_2 D_2^{-1} \right] \begin{bmatrix} {}^t X_1 \\ {}^t X_2 \end{bmatrix} \\ &= \frac{1}{2n} \left[X_1 D_1^{-1} {}^t X_1 + X_2 D_2^{-1} {}^t X_2 \right] . \end{aligned}$$

Cette matrice est $n \times n$ et l'on en cherche les vecteurs propres nI_n -orthonormés ($\bar{D}_1^{-1} = n I_n$) ; on sait que seuls les $\ell-1$ vecteurs propres associés aux $\ell-1$ plus grandes valeurs propres autres que 1 ($\frac{1+\sqrt{\lambda_k}}{2}$; $k=1, \dots, \ell-1$) sont intéressants ; soit \bar{U}_1 la matrice $n \times (\ell-1)$ de ces vecteurs propres $n I_n$ -orthonormés disposés en colonnes ; on a :

$$\begin{aligned}\bar{U}_1 &= {}^t\bar{A}_2 \bar{U}_2 N^{-1/2} = \frac{1}{2n} [X_1 \bar{D}_1^{-1} \mid X_2 \bar{D}_2^{-1}] \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} N^{-1/2} \\ &= \frac{1}{2n} [X_1 \bar{D}_1^{-1} U_1 + X_2 \bar{D}_2^{-1} U_2] N^{-1/2} ,\end{aligned}$$

avec $N = \text{diag}(v_1, \dots, v_{\ell-1}) = \frac{1}{2} [I_{\ell-1} + L^{1/2}]$.

• Matrice des composantes principales :

$$\begin{aligned}\text{c'est } \bar{C}_2 &= \bar{A}_2 \bar{D}_1^{-1} \bar{U}_1 = \begin{bmatrix} \bar{D}_1^{-1} {}^tX_1 \\ \bar{D}_2^{-1} {}^tX_2 \end{bmatrix} \bar{U}_1 \\ &= \frac{1}{2n} \begin{bmatrix} \bar{D}_1^{-1} {}^tX_1 X_1 \bar{D}_1^{-1} U_1 + \bar{D}_1^{-1} {}^tX_1 X_2 \bar{D}_2^{-1} U_2 \\ \bar{D}_2^{-1} {}^tX_2 X_1 \bar{D}_1^{-1} U_1 + \bar{D}_2^{-1} {}^tX_2 X_2 \bar{D}_2^{-1} U_2 \end{bmatrix} N^{-1/2} \\ &= \frac{1}{2} \begin{bmatrix} \bar{D}_1^{-1} U_1 + \bar{D}_1^{-1} {}^tA_2 U_2 \\ \bar{D}_2^{-1} {}^tA_1 U_1 + \bar{D}_2^{-1} U_2 \end{bmatrix} N^{-1/2} = \frac{1}{2} \begin{bmatrix} \bar{D}_1^{-1} U_1 (I_{\ell-1} + L^{1/2}) \\ \bar{D}_2^{-1} U_2 (L^{1/2} + I_{\ell-1}) \end{bmatrix} N^{-1/2}\end{aligned}$$

(pour cette dernière égalité, on utilise la propriété $U_1 = {}^tA_2 U_2 L^{-1/2} \Rightarrow {}^tA_2 U_2 = U_1 L^{1/2}$; pareillement ${}^tA_1 U_1 = U_2 L^{1/2}$).

$$\text{On obtient ainsi : } \bar{C}_2 = \frac{1}{2} \begin{bmatrix} \bar{D}_1^{-1} U_1 \\ \bar{D}_2^{-1} U_2 \end{bmatrix} (I_{\ell-1} + L^{1/2}) N^{-1/2} = \begin{bmatrix} \bar{D}_1^{-1} U_1 \\ \bar{D}_2^{-1} U_2 \end{bmatrix} N^{1/2} .$$

En utilisant alors la remarque du 2.5., il vient finalement :

$$\bar{C}_2 = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} L^{-1/2} N^{1/2} .$$

• Représentation des modalités de X^1 et de X^2 :

L'A.C.P. des profils-colonnes de X a pour objectif de représenter simultanément les modalités de X^1 et celles de X^2 au moyen de la matrice \bar{C}_2 ; on constate que cela revient à représenter les modalités de X^1 au moyen de $C_1 L^{-1/2} N^{1/2}$ et celles de X^2 au moyen de $C_2 L^{-1/2} N^{1/2}$; cette représentation est voisine de la représentation simultanée brute de l'A.F.C. classique, mais elle ne lui est pas identique, la coordonnée sur chaque axe (c_i^k ou c_j^k) étant ici multipliée par le facteur $\sqrt{\frac{1+\sqrt{\lambda_k}}{2\lambda_k}}$; pour fixer les idées, voici les valeurs de ce facteur pour quelques valeurs typiques de λ_k (toujours comprises entre 1 et 0) :

valeurs de λ_k	valeurs de $\sqrt{\frac{1+\sqrt{\lambda_k}}{2\lambda_k}}$
0,90	1,04
0,75	1,12
0,50	1,31
0,25	1,73
0,10	2,57

ce n'est donc que pour des valeurs très distinctes de λ_k (ce qui est rare puisqu'on s'intéresse essentiellement à λ_1 , λ_2 et λ_3), que l'on peut avoir une modification sensible de la représentation graphique.

• Représentation des individus.

L'A.C.P. des profils-lignes de X a, elle, pour objectif de représenter les individus étudiés au moyen de la matrice \bar{C}_1 .

On a obtenu : $\bar{C}_1 = \frac{1}{2} [X^1 D_1^{-1} U_1 + X^2 D_2^{-1} U_2]$; compte-tenu de la forme des matrices X^1 et X^2 , cela signifie qu'un individu ayant présenté les modalités i de X^1 et j de X^2 sera représenté sur le k-ième axe par la valeur $\frac{1}{2} [(D_1^{-1} U_1)_i^k + (D_2^{-1} U_2)_j^k]$.

Il s'agirait de la moyenne vectorielle entre les représentations de ces deux modalités si leurs coordonnées sur le k-ième axe étaient $(D_1^{-1} U_1)_i^k$ et $(D_2^{-1} U_2)_j^k$; nous avons vu qu'en réalité, ces valeurs sont multipliées par le facteur

$\sqrt{\frac{1+\sqrt{\lambda_k}}{2}}$ sur chaque axe ; toujours pour les mêmes valeurs typiques de λ_k , ce facteur vaut :

valeurs de λ_k	valeurs de $\sqrt{\frac{1+\sqrt{\lambda_k}}{2}}$
0,90	0,99
0,75	0,97
0,50	0,92
0,25	0,87
0,10	0,81

On peut donc en déduire qu'en général, la représentation de chaque individu correspondra approximativement à la moyenne vectorielle des représentations des modalités de chaque variable prises par cet individu.

3.5. Quelques remarques complémentaires.

- Mesure de la qualité des représentations graphiques :

Dans cette analyse, à cause des valeurs propres non nulles sans signification, les indicateurs de qualité de représentation de chaque individu ou de chaque modalité de X^1 ou de X^2 (cosinus d'angle) n'ont plus de signification réelle (c'est le même problème que pour l'indicateur de qualité globale \bar{r}_k).

- Signification des contributions relatives :

Les contributions relatives à l'inertie de chaque nuage ou à la détermination des différents axes principaux conservent ici la même signification que dans une A.F.C. simple ; toutefois, il faut noter que d'une part on obtient ces contributions pour les individus, ce qui n'est pas le cas dans une A.F.C. simple, et que d'autre part les contributions des modalités de X^1 et de X^2 sont données globalement et non pas séparément, pour chaque variable. Signalons enfin que les contributions des modalités à l'inertie de leur nuage ne sont pas comparables à celles obtenues dans l'A.F.C. simple (les espaces ne sont pas les mêmes) alors que les contributions des modalités à la détermination des axes sont divisées par deux (pour chaque variable, l'inertie de ses modalités selon le k-ième axe vaut λ_k dans l'A.F.C. simple, et les coordonnées de ces modalités subissent ici une homothétie sur chaque axe).

3.6. Exemple:

Nous reprenons le même exemple qu'en 2.7. ; le tableau disjonctif complet, avec ses marges, est le suivant :

	X_1^1	X_2^1	X_3^1	X_1^2	X_2^2	X_3^2	X_4^2	
1	0	1	0	0	1	0	0	2
2	0	0	1	0	0	1	0	2
3	0	0	1	0	0	0	1	2
4	1	0	0	1	0	0	0	2
5	0	1	0	0	0	1	0	2
6	1	0	0	0	1	0	0	2
	2	2	2	1	2	2	1	12

On peut vérifier que l'on a :

$${}^t\bar{A}_1 \quad {}^t\bar{A}_2 = \frac{1}{4} \times \begin{bmatrix} 2 & 0 & 0 & \vdots & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & \vdots & 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & \vdots & 0 & 0 & 1 & 1 \\ \hline 1 & 0 & 0 & \vdots & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & \vdots & 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & \vdots & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & \vdots & 0 & 0 & 0 & 1 \end{bmatrix} ;$$

cette matrice admet pour valeurs propres :

$$v_0 = 1 \quad v_1 \approx 0,933 \quad v_2 = 0,75 \quad v_3 = 0,50 \quad v_4 = 0,25 \quad v_5 \approx 0,067 \quad v_6 = 0 ;$$

seules v_1 et v_2 sont utiles ici.

De la même manière, on a :

$${}^t\bar{A}_2 \quad {}^t\bar{A}_1 = \frac{1}{4} \times \begin{bmatrix} 2 & 0 & 0 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 & 0 & 2 \end{bmatrix} ;$$

les valeurs propres sont :

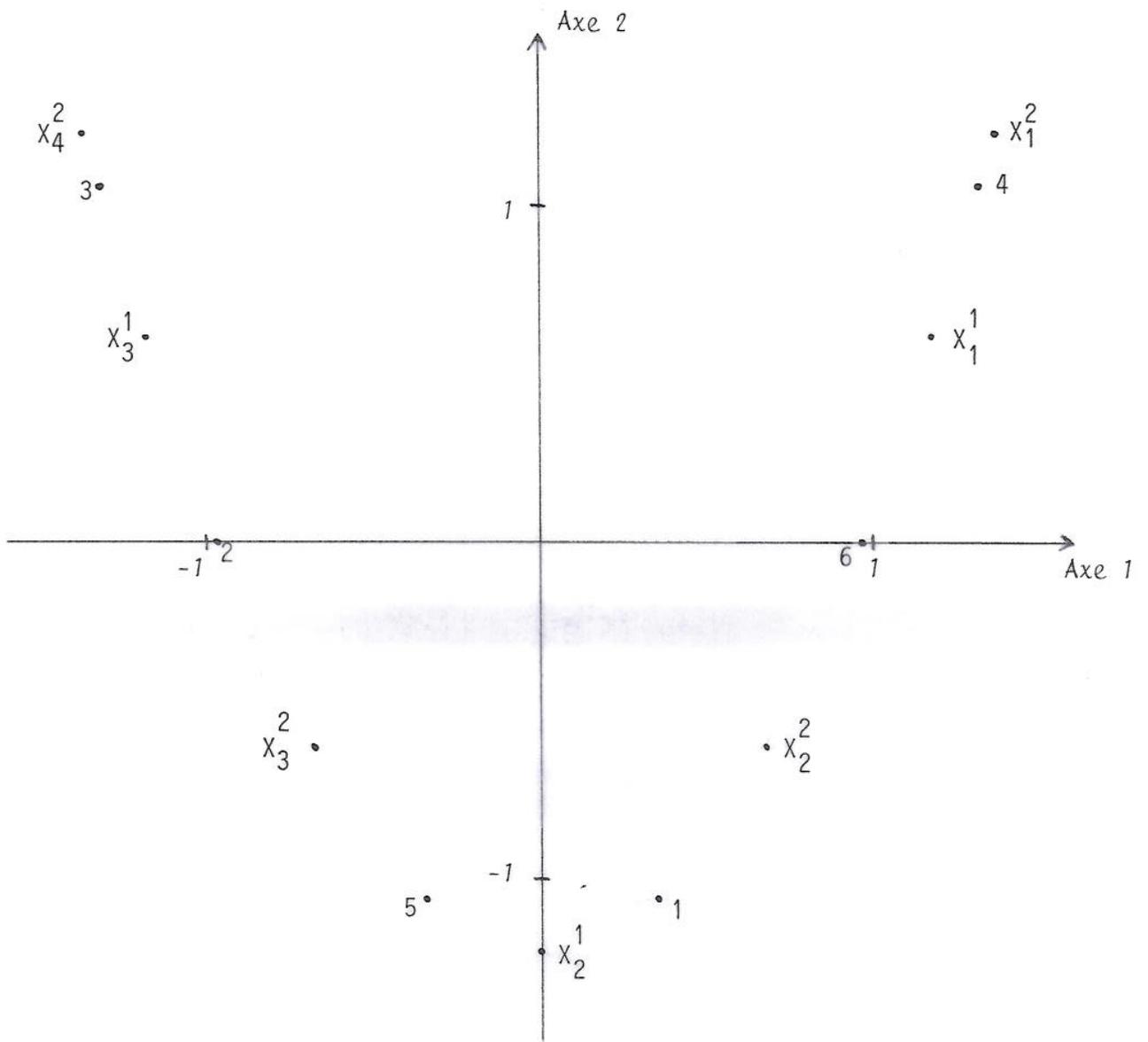
$$v_0 = 1 \quad v_1 \approx 0,933 \quad v_2 = 0,75 \quad v_3 = 0,50 \quad v_4 = 0,25 \quad v_5 \approx 0,067 .$$

On vérifie bien que $\sum_{k=1}^5 v_k = 2,50$.

On obtient ensuite :

$$\bar{C}_1 \approx \begin{bmatrix} 0,35 & -1,06 \\ -0,97 & 0 \\ -1,32 & 1,06 \\ 1,32 & 1,06 \\ -0,35 & -1,06 \\ 0,97 & 0 \end{bmatrix} ; \quad \bar{C}_2 \approx \begin{bmatrix} 1,18 & 0,61 \\ 0 & -1,22 \\ -1,18 & 0,61 \\ \hline 1,37 & 1,22 \\ 0,68 & -0,61 \\ -0,68 & -0,61 \\ -1,37 & 1,22 \end{bmatrix} .$$

\bar{C}_2 permet de faire une représentation simultanée brute de l'ensemble des modalités de X^1 et de X^2 , peu différente de celle obtenue en 2.7. à partir de C_1 et C_2 ; \bar{C}_1 permet de représenter les individus ; la technique d'A.F.C. permet de représenter l'ensemble sur un graphique unique dans lequel chaque individu est proche du barycentre des modalités qu'il a présentées. Ce graphique est ici le suivant :



Les contributions relatives de l'ensemble des modalités à l'inertie de leur nuage sont :

$$\begin{array}{cccc} X_1^1 : 13\% & X_2^1 : 13\% & X_3^1 : 13\% & \\ X_1^2 : 17\% & X_2^2 : 13\% & X_3^2 : 13\% & X_4^2 : 17\% \end{array}$$

Pour les individus, on obtient :

$$1 : 13\% \quad 2 : 13\% \quad 3 : 23\% \quad 4 : 23\% \quad 5 : 13\% \quad 6 : 13\%.$$

Les contributions relatives des modalités à la détermination des axes sont :

$$\begin{array}{l} \text{pour l'axe 1 ,} \\ X_1^1 : 25\% \quad X_2^1 : 0\% \quad X_3^1 : 25\% \\ X_1^2 : 17\% \quad X_2^2 : 8\% \quad X_3^2 : 8\% \quad X_4^2 : 17\% ; \end{array}$$

$$\begin{array}{l} \text{pour l'axe 2 ,} \\ X_1^1 : 8\% \quad X_2^1 : 33\% \quad X_3^1 : 8\% \\ X_1^2 : 17\% \quad X_2^2 : 8\% \quad X_3^2 : 8\% \quad X_4^2 : 17\% ; \end{array}$$

Les contributions relatives des individus à la détermination des axes sont :

$$\begin{array}{l} \text{pour l'axe 1 ,} \quad 1 : 2\% \quad 2 : 17\% \quad 3 : 31\% \quad 4 : 31\% \quad 5 : 2\% \quad 6 : 17\% ; \\ \text{pour l'axe 2 ,} \quad 1 : 25\% \quad 2 : 0\% \quad 3 : 25\% \quad 4 : 25\% \quad 5 : 25\% \quad 6 : 0\% . \end{array}$$

4. A.F.C. DU TABLEAU DE BURT RELATIF A DEUX VARIABLES QUALITATIVES

4.1. Définition du tableau de BURT.

On dispose toujours de deux variables qualitatives X^1 , à l modalités, et X^2 , à c modalités, observées simultanément sur n individus affectés du même poids ; les notations seront encore les mêmes que celles introduites en 2. et en 3..

On appelle tableau de BURT relatif à X^1 , à X^2 , et aux n individus observés, le tableau carré symétrique $B = {}^tXX$.

Comme on a $X = [X_1 \mid X_2]$, il vient :

$$B = \left[\begin{array}{c|c} t_{X_1 X_1} & t_{X_1 X_2} \\ \hline t_{X_2 X_1} & t_{X_2 X_2} \end{array} \right] = \left[\begin{array}{c|c} n D_1 & T \\ \hline t_T & n D_2 \end{array} \right] .$$

Les effectifs marginaux des lignes (identiques aux effectifs marginaux des colonnes) valent $2n_{1.}, \dots, 2n_{\ell.}, 2_{n.1}, \dots, 2_{n.c}$, et l'effectif total est égal à $4n$.

4.2. Principe et notations.

Le principe est ici de faire l'A.F.C. du tableau B ; compte-tenu que B est symétrique, une seule A.C.P. sera nécessaire.

Les quantités intervenant dans cette A.F.C. et calculées à partir de B seront toujours notées de la même manière, mais surmontées cette fois du signe \sim .

Il vient ainsi :

$$\tilde{T} = B ; \quad \tilde{D}_1 = \frac{1}{2} \left[\begin{array}{c|c} D_1 & 0 \\ \hline 0 & D_2 \end{array} \right] = \tilde{D}_2 \quad ; \quad \tilde{D}_1^{-1} = \tilde{D}_2^{-1} = 2 \left[\begin{array}{c|c} D_1^{-1} & 0 \\ \hline 0 & D_2^{-1} \end{array} \right] ;$$

$$\tilde{A}_1 = \frac{1}{2} \left[\begin{array}{c|c} I_\ell & A_1 \\ \hline A_2 & I_c \end{array} \right] = \tilde{A}_2 .$$

4.3. A.C.P. des profils-lignes de B.

• Matrice à diagonaliser et matrice des vecteurs propres : La matrice à diagonaliser est ${}^t\tilde{A}_1 \tilde{A}_2 = [{}^t\tilde{A}_1]^2 = [{}^t\tilde{A}_1 \tilde{A}_2]^2$: on doit diagonaliser le carré de la matrice diagonalisée en 3.3. ; on sait qu'une matrice et son carré ont les mêmes vecteurs propres, les valeurs propres étant élevées au carré ; comme de plus, $\tilde{D}_2^{-1} = \overline{D}_2^{-1}$, on conclut immédiatement que ${}^t\tilde{A}_1 \tilde{A}_2$ n'a que $(\ell-1)$ vecteurs propres intéressants pour cette analyse ; si on les dispose en colonnes dans la matrice $\tilde{U}_2, (\ell+c) \times (\ell-1)$, il vient :

$$[\tilde{U}_2] = [\overline{U}_2] = \frac{1}{2} \left[\begin{array}{c} U_1 \\ U_2 \end{array} \right] .$$

Par ailleurs, les $(\ell-1)$ valeurs propres utiles rangées par ordre décroissant s'écrivent :

$$\rho_k = v_k^2 = \frac{[1 + \sqrt{\lambda_k}]^2}{4} , \quad k = 1, \dots, \ell-1 ;$$

Remarque : Là encore, l'indicateur de qualité $\tilde{r}_k = \frac{\rho_k}{\sum_{k=1}^{\ell+c-1} \rho_k}$ n'a pas de signification réelle.

Toutefois, en remarquant que $\sum_{k=1}^{\ell+c-1} \rho_k = \frac{\ell+c+2(\phi^2-1)}{4}$, on pourrait dans ce cas

calculer :
$$\frac{[\sqrt{[\ell+c+2(\phi^2-1)]r_k} - 1]^2}{\phi^2} .$$

• Matrice des composantes principales :

Elle s'écrit :
$$\tilde{C}_1 = \tilde{A}_1 \tilde{D}_2^{-1} \tilde{U}_2 = \frac{1}{2} \left[\begin{array}{c|c} I_\ell & A_1 \\ \hline A_2 & I_c \end{array} \right] \left[\begin{array}{c|c} D_1^{-1} & 0 \\ \hline 0 & D_2^{-1} \end{array} \right] \left[\begin{array}{c} U_1 \\ U_2 \end{array} \right]$$

$$= \frac{1}{2} \left[\begin{array}{c|c} D_1^{-1} & A_1 D_2^{-1} \\ \hline A_2 D_1^{-1} & D_2^{-1} \end{array} \right] \left[\begin{array}{c} U_1 \\ U_2 \end{array} \right] = \frac{1}{2} \left[\begin{array}{c} D_1^{-1} U_1 + A_1 D_2^{-1} U_2 \\ \hline A_2 D_1^{-1} U_1 + D_2^{-1} U_2 \end{array} \right]$$

$$= \frac{1}{2} \left[\begin{array}{c} D_1^{-1} U_1 + D_1^{-1} {}^t A_2 U_2 \\ \hline D_2^{-1} A_1 U_1 + D_2^{-1} U_2 \end{array} \right] = \frac{1}{2} \left[\begin{array}{c} D_1^{-1} U_1 \\ \hline D_2^{-1} U_2 \end{array} \right] (I_{\ell-1} + L^{1/2})$$

(d'après 3.4)

$$= \left[\begin{array}{c} D_1^{-1} U_1 \\ \hline D_2^{-1} U_2 \end{array} \right] N ; \quad \text{on a finalement : } \tilde{C}_1 = \left[\begin{array}{c} C_1 \\ \hline C_2 \end{array} \right] L^{-1/2} N.$$

Comme on a dans ce cas $\tilde{C}_2 = \tilde{C}_1$, il n'y a pas à faire de représentation simultanée, \tilde{C}_1 permettant de représenter les modalités de X^1 et celles de X^2 sur un seul graphique.

Sur ce graphique, la représentation des modalités de X^1 et de X^2 est donc encore quelque peu différente de ce qu'elle est dans les autres cas ; par rapport à l'A.F.C. classique (celle de T), les coordonnées sur le k-ième axe sont ainsi multipliées par le facteur $\frac{1}{2} \left(1 + \frac{1}{\sqrt{\lambda_k}}\right)$ dont nous donnons encore les valeurs dans les mêmes cas que précédemment :

valeur de λ_k	valeur de $\frac{1}{2} \left(1 + \frac{1}{\sqrt{\lambda_k}} \right)$
0,90	1,03
0,75	1,08
0,50	1,21
0,25	1,50
0,10	2,08

On constate donc que la déformation est moindre dans ce cas que dans le cas de l'analyse du tableau disjonctif complet.

4.4. Remarques complémentaires.

- Inutilité de l'A.C.P. des profils-colonnes :

Il va de soi que l'A.C.P. des profils-colonnes de B conduit aux mêmes résultats que la précédente, B étant symétrique ; il est donc inutile de la développer.

- Mesure de la qualité des représentations graphiques :

Pour les mêmes raisons qu'au point 3. (valeurs propres non nulles sans signification réelle), la qualité de représentation de chaque modalité de X^1 ou de X^2 (cosinus d'angle) n'a pas de signification, comme l'indicateur de qualité globale \tilde{r}_k .

- Contributions relatives :

Les contributions relatives des modalités de X^1 et de X^2 à l'inertie totale de leur nuage sont, comme dans le cas précédent, données globalement ; toutefois, elles ne sont pas comparables à celles obtenues au paragraphe 3. puisque l'espace n'est plus le même ($\mathbb{R}^{\ell+c}$ au lieu de \mathbb{R}^n), ni bien sûr à celles obtenues au paragraphe 2., de façons séparées.

Les contributions relatives des modalités de chaque variable à la détermination des axes sont identiques à celles obtenues dans l'A.F.C. du tableau disjonctif complet et donc égales à la moitié de celles obtenues dans l'A.F.C. de la table de contingence (pour les mêmes raisons que celles exposées en 3.5.).

Enfin, remarquons que l'on ne dispose plus ici de contributions pour les individus, ces derniers n'étant pas représentés dans cette analyse.

4.5. Exemple.

Nous reprenons le même exemple qu'en 2.7. et 3.6. ; le tableau de BURT, avec ses marges, est le suivant :

B :

	x_1^1	x_2^1	x_3^1	\vdots	x_1^2	x_2^2	x_3^2	x_4^2	
x_1^1	2	0	0	\vdots	1	1	0	0	4
x_2^1	0	2	0	\vdots	0	1	1	0	4
x_3^1	0	0	2	\vdots	0	0	1	1	4
<hr style="border-top: 1px dashed black;"/>									
x_1^2	1	0	0	\vdots	1	0	0	0	2
x_2^2	1	1	0	\vdots	0	2	0	0	4
x_3^2	0	1	1	\vdots	0	0	2	0	4
x_4^2	0	0	1	\vdots	0	0	0	1	2
	4	4	4	\vdots	2	4	4	2	24

La matrice à diagonaliser est le carré de la matrice ${}^t\bar{A}_1 {}^t\bar{A}_2$ écrite en 3.6. (elle est 7×7) ; ses valeurs propres sont :

$$\rho_0 = 1 \quad \rho_1 \approx 0,871 \quad \rho_2 \approx 0,563 \quad \rho_3 = 0,25 \quad \rho_4 \approx 0,063$$

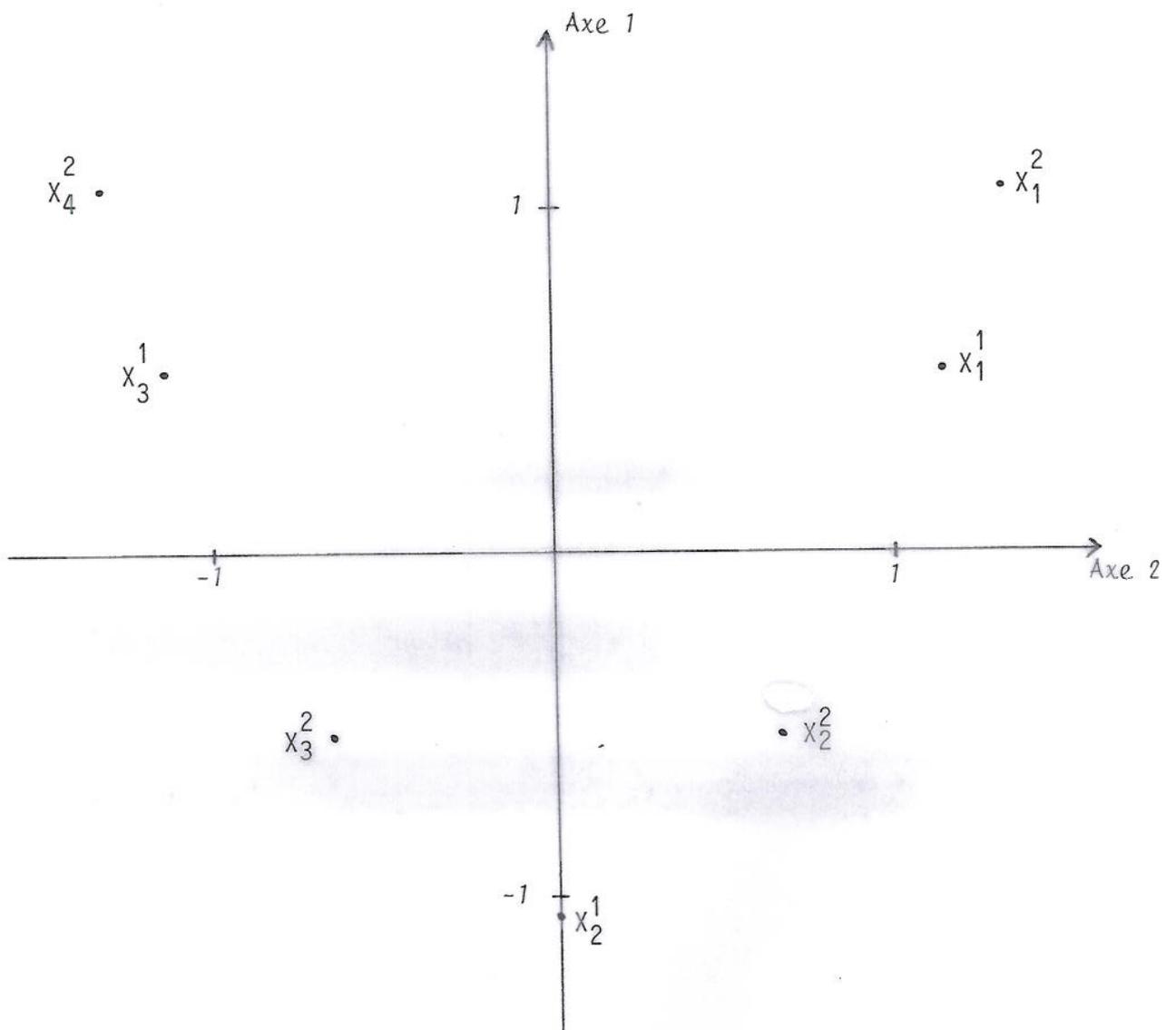
$\rho_5 \approx 0,004 \quad \rho_6 = 0$; encore une fois, seules ρ_1 et ρ_2 sont utiles.

On pourra vérifier que $\sum_{k=1}^{\lambda+c-1} \rho_k = \frac{\lambda+c+2(\phi^2-1)}{4} = 1,75$ (on a vu en 2.7. que $\phi^2 = 1$).

Les matrices des composantes principales sont :

$$\tilde{C}_1 = \tilde{C}_2 \approx \begin{bmatrix} 1,14 & 0,53 \\ 0 & -1,06 \\ -1,14 & 0,53 \\ \hline 1,32 & 1,06 \\ 0,66 & -0,53 \\ -0,66 & -0,53 \\ -1,32 & 1,06 \end{bmatrix} .$$

La représentation graphique des modalités de X^1 et de X^2 réalisée à partir de \tilde{C}_1 (voir ci-dessous) est encore peu différente des deux précédentes.



Les contributions relatives de l'ensemble des modalités à l'inertie de leur nuage valent ici :

$$\begin{array}{llll} X_1^1 : 15\% & X_2^1 : 12\% & X_3^1 : 15\% & \\ X_1^2 : 17\% & X_2^2 : 12\% & X_3^2 : 12\% & X_4^2 : 17\% . \end{array}$$

Pour les contributions relatives des modalités à la détermination des axes, on se reportera au paragraphe 3.6., les résultats étant identiques.

5. A.F.C. D'UN TABLEAU DISJONCTIF COMPLET RELATIF A PLUS DE DEUX VARIABLES QUALITATIVES

5.1. Données et notations.

Nous considérons ici p variables qualitatives ($p \geq 3$) $X^1, \dots, X^j, \dots, X^p$ observées simultanément sur n individus affectés du même poids $\frac{1}{n}$. Nous noterons c_j le nombre de modalités de X^j ($j=1, \dots, p$) et $c = \sum_{j=1}^p c_j$.

Soit X le tableau disjonctif complet $n \times c$ relatif à ces observations ; si l'on note X_j le tableau des indicatrices $n \times c_j$ associé à X^j , on a :

$$X = [X_1 \mid \dots \mid X_j \mid \dots \mid X_p].$$

Soit n_i^j l'effectif de la i -ième modalité de la j -ième variable ($1 \leq j \leq p$; $1 \leq i \leq c_j$) ; nous poserons $D_j = \frac{1}{n} \text{diag}(n_1^j, \dots, n_{c_j}^j)$ et $\Delta = \begin{bmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_p \end{bmatrix}$.

Les effectifs marginaux des lignes de X valent tous p , les effectifs marginaux des colonnes valent $n_1^1, \dots, n_{c_1}^1, \dots, n_1^p, \dots, n_{c_p}^p$ et l'effectif total vaut np .

\bar{D}_1 , matrice $n \times n$, diagonale, des fréquences des lignes vaut $\frac{1}{n} I_n$; d'où $\bar{D}_1^{-1} = n I_n$.

\bar{D}_2 , matrice $c \times c$, diagonale, des fréquences des colonnes vaut $\frac{1}{p} \Delta$;

$$\text{d'où : } \bar{D}_2^{-1} = p \Delta^{-1}.$$

Soit \bar{A}_1 la matrice $n \times c$ des profils-lignes associés à X et disposés en lignes ; on a : $\bar{A}_1 = \frac{1}{p} X$.

\bar{A}_2 , matrice $c \times n$ des profils-colonnes disposés en lignes, s'écrit : $\bar{A}_2 = \frac{1}{n} \Delta^{-1} t_X$.

5.2. A.C.P. des profils-lignes de X.

• Matrice à diagonaliser :

C'est $t_{\bar{A}_1} t_{\bar{A}_2} = \frac{1}{np} t_{XX} \Delta^{-1} = \frac{1}{np} B \Delta^{-1}$, où $B = t_{XX}$ est le tableau de BURT ($c \times c$ et symétrique) relatif aux p variables X^1, \dots, X^p .

Posons $S = \frac{1}{n} B \Delta^{-1}$; S est une matrice $c \times c$ constituée de p^2 blocs, le bloc courant s'écrivant $\frac{1}{n} {}^t X_j X_g D_g^{-1}$ et représentant la matrice des profils-colonnes disposés en colonnes pour le croisement des deux variables X^j et X^g prises dans cet ordre (pour ce croisement, c'est la matrice ${}^t A_2$) ; S généralise la matrice

matrice $\left[\begin{array}{c|c} I_\ell & {}^t A_2 \\ \hline {}^t A_1 & I_c \end{array} \right]$ étudiée en 3.3.. On peut remarquer que ${}^t S = \frac{1}{n} \Delta^{-1} B$.

• Matrice des vecteurs propres :

D'après les résultats généraux de l'A.F.C., on sait que ${}^t \bar{A}_1 {}^t \bar{A}_2 = \frac{1}{p} S$ admet c valeurs propres comprises entre 0 et 1, la plus grande valant 1 et n'ayant pas d'intérêt dans l'analyse ; d'autre part, $(p-1)$ de ces valeurs propres sont nulles, compte-tenu que les variables intervenant en colonnes dans X sont des indicatrices liées par p liaisons linéaires ; soit donc m le nombre de valeurs propres strictement positives en dehors de 1 ($m \leq c-p$) et soient v_1, \dots, v_m ces valeurs propres rangées par ordre décroissant ; appelons u_1, \dots, u_m les vecteurs propres \bar{D}_2^{-1} -orthonormés (c'est-à-dire $p \Delta^{-1}$ -orthonormés) associés, et \bar{U}_2 la matrice $c \times m$ de ces vecteurs disposés en colonnes.

$$\forall k=1, \dots, m, \text{ on a } [u^k] = \begin{bmatrix} u_1^k \\ \vdots \\ u_p^k \end{bmatrix}, \text{ où } u_j^k \in \mathbb{R}^{c_j} ;$$

$$\text{on a de même } \bar{U}_2 = \begin{bmatrix} U_1 \\ \vdots \\ U_p \end{bmatrix}, \text{ où } U_j = [u_j^1 | \dots | u_j^m] \\ (U_j \text{ est } c_j \times m).$$

Notant enfin $N = \text{diag}(v_1, \dots, v_m)$, il vient :

$$S \bar{U}_2 = p \bar{U}_2 N \quad (3).$$

Remarque : Contrairement au cas $p=2$, il n'est pas possible ici d'imposer des conditions de normalisation aux u_j^k , notamment parce qu'ils ne sont pas vecteurs propres de matrices connues.

On notera également qu'il n'y a pas ici de relation entre les différentes matrices U_j ($j=1, \dots, p$).

• Matrice des composantes principales :

Elle s'écrit : $\bar{C}_1 = \bar{A}_1 \bar{D}_2^{-1} \bar{U}_2 = X \Delta^{-1} \bar{U}_2$; il vient : $\bar{C}_1 = \sum_{j=1}^p X_j D_j^{-1} U_j$.

Nous utiliserons ce résultat en 5.3..

5.3. A.C.P. des profils-colonnes de X.

• Matrice à diagonaliser et matrice des vecteurs propres :

On diagonalise ici ${}^t \bar{A}_2 \bar{A}_1 = \frac{1}{np} X \Delta^{-1} X$

$$= \frac{1}{np} \sum_{j=1}^p X_j D_j^{-1} X_j .$$

Ses valeurs propres non nulles sont les mêmes que celles de ${}^t \bar{A}_1 \bar{A}_2$, soient v_1, \dots, v_m ; soit U_1 la matrice $n \times m$ donnant dans ses colonnes les vecteurs propres \bar{D}_1^{-1} -orthonormés (c'est-à-dire $n I_n$ -orthonormés) de ${}^t \bar{A}_2 \bar{A}_1$; on a :

$$\bar{U}_1 = {}^t \bar{A}_2 \bar{U}_2 N^{-1/2} = \frac{1}{n} X \Delta^{-1} \bar{U}_2 N^{-1/2} .$$

• Matrice des composantes principales :

C'est : $\bar{C}_2 = \bar{A}_2 \bar{D}_1^{-1} \bar{U}_1 = \frac{1}{n} \Delta^{-1} X X \Delta^{-1} \bar{U}_2 N^{-1/2} = \Delta^{-1} S \bar{U}_2 N^{-1/2}$

$$= p \Delta^{-1} \bar{U}_2 N^{1/2} \text{ (d'après (3))}$$

$$= p \begin{bmatrix} D_1^{-1} U_1 \\ \vdots \\ D_p^{-1} U_p \end{bmatrix} N^{1/2} ;$$

On retrouve ici une formule analogue à celle du paragraphe 3.4. ; en particulier, si l'on pose $C_j = p D_j^{-1} U_j N^{1/2}$ ($j=1, \dots, p$), on a alors :

$$\bar{C}_2 = \begin{bmatrix} C_1 \\ \vdots \\ C_p \end{bmatrix} .$$

On notera toutefois qu'ici U_j n'est pas nécessairement une matrice de vecteurs colonnes D_j^{-1} -orthonormés. En particulier, le coefficient p généralise le coefficient 2 implicitement apparu au paragraphe 3.4. dans l'expression de \bar{C}_2 , ce coefficient s'étant ensuite simplifié avec le facteur $\frac{1}{2}$ introduit pour retrouver en U_1 (respectivement U_2) une matrice de vecteurs D_1^{-1} (respectivement D_2^{-1})-orthonormés.

- Représentation des modalités des variables X^j :

Chaque modalité i de chaque variable X^j est représentée au moyen de la i -ième ligne de $C_j = p D_j^{-1} U_j N^{1/2}$.

- Représentation des individus :

Elle se fait au moyen des lignes de $\bar{C}_1 = \frac{p}{\sum_{j=1}^p X_j} D_j^{-1} U_j = \frac{1}{p} \frac{p}{\sum_{j=1}^p X_j} C_j N^{-1/2}$;

donc, là encore, chaque individu est représenté par le vecteur moyen associé à l'ensemble des vecteurs représentant les modalités des variables X^j que cet individu a présentés ; toutefois, la représentation des modalités de chaque variable X^j doit pour cela se faire au moyen des lignes de la matrice $C_j N^{-1/2}$ qui est quelque peu différente de la matrice C_j .

5.4. Bilan.

- Graphique obtenu :

La technique d'A.F.C. (et notamment la représentation simultanée brute) appliquée au tableau disjonctif complet relatif à p variables qualitatives ($p \geq 3$) permet de représenter sur un même graphique les modalités de l'ensemble des variables considérées (matrice \bar{C}_2) ainsi que les individus (matrice \bar{C}_1), chacun d'eux étant proche du barycentre des vecteurs associés aux modalités qu'il a présentées.

Cette technique constitue une généralisation naturelle de celle étudiée dans le cas $p=2$, cette dernière étant sensiblement identique à l'A.F.C. classique de la table de contingence T ; elle constitue la *première façon de pratiquer l'Analyse Factorielle des Correspondances Multiple*.

- Autres résultats :

En plus de la représentation graphique décrite ci-dessus, les logiciels d'A.F.C. appliqués au tableau disjonctif complet X fournissent des indicateurs de qualité et des contributions relatives.

Comme dans le cas $p=2$, et pour les mêmes raisons, les indicateurs de qualité (qualité globale en pourcentage d'inertie et qualité de la représentation des individus et des modalités en cosinus d'angles) n'ont pas de signification concrète et ne sont en général pas utilisés. Les diverses contributions relatives sont, elles, utiles à l'interprétation du graphique et ont la même signification que celle indiquée en 3.3. .

5.5. Exemple.

Nous reprenons ici l'exemple utilisé précédemment et nous considérons en plus une troisième variable qualitative X^3 à deux modalités notées X_1^3 et X_2^3 ; le tableau disjonctif complet, avec ses marges, est le suivant :

	X_1^1	X_2^1	X_3^1	\vdots	X_1^2	X_2^2	X_3^2	X_4^2	\vdots	X_1^3	X_2^3	
1	0	1	0	\vdots	0	1	0	0	\vdots	1	0	3
2	0	0	1	\vdots	0	0	1	0	\vdots	0	1	3
3	0	0	1	\vdots	0	0	0	1	\vdots	0	1	3
4	1	0	0	\vdots	1	0	0	0	\vdots	1	0	3
5	0	1	0	\vdots	0	0	1	0	\vdots	1	0	3
6	1	0	0	\vdots	0	1	0	0	\vdots	1	0	3
	2	2	2	\vdots	1	2	2	1	\vdots	4	2	18

$t_{\bar{A}_1} \quad t_{\bar{A}_2} = \frac{1}{np} B \Delta^{-1} = \frac{1}{3} B \Gamma^{-1}$, où B est le tableau de BURT associé à X^1 , X^2 et X^3 ($B = {}^t X X$, est explicité au paragraphe 6.4.) et où $\Gamma = \text{diag}(2,2,2, 1,2,2,1,4,2)$.

Les valeurs propres sont :

$$v_0 = 1 \quad v_1 \approx 0,913 \quad v_2 \approx 0,520 \quad v_3 \approx 0,333 \quad v_4 \approx 0,180$$

$$v_5 \approx 0,054 \quad v_6 = v_7 = v_8 = 0 \text{ (on a ici } m=5 \text{).}$$

$$t_{\bar{A}_2} \quad t_{\bar{A}_1} = \frac{1}{np} \sum_{j=1}^p X_j D_j^{-1} t_{X_j} = \frac{1}{3} \sum_{j=1}^3 X_j \Gamma_j^{-1} t_{X_j} ,$$

où $\Gamma_1 = \text{diag}(2,2,2)$, $\Gamma_2 = \text{diag}(1,2,2,1)$ et $\Gamma_3 = \text{diag}(4,2)$.

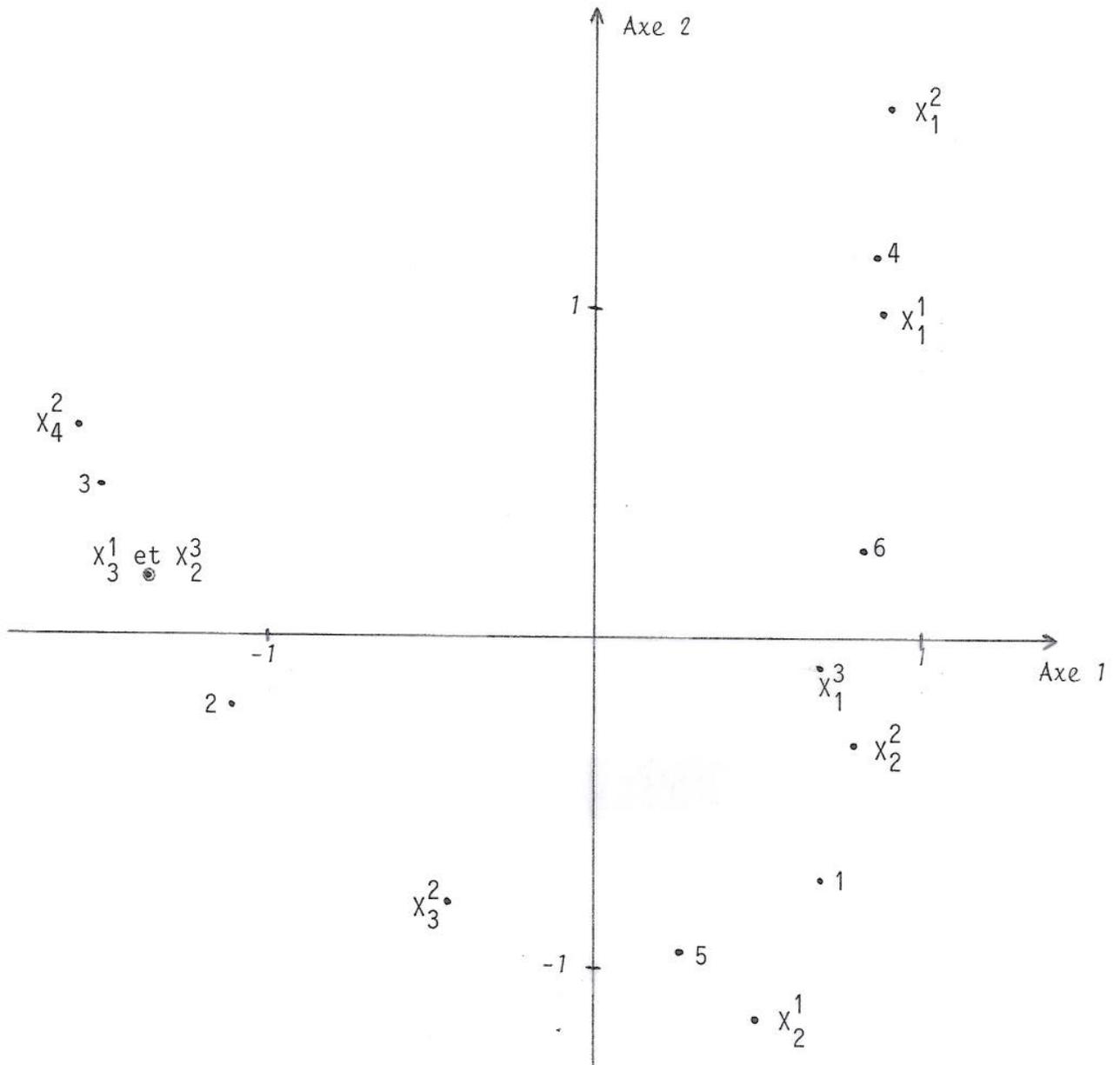
Les valeurs propres de $t_{\bar{A}_2} \quad t_{\bar{A}_1}$ sont v_0, v_1, v_2, v_3, v_4 et v_5 .

Les matrices des composantes principales \bar{C}_1 et \bar{C}_2 ont ici 5 colonnes chacune ; pour réaliser une représentation graphique en dimension 2, on ne considère que les 2 premières colonnes de chacune d'elles ; en notant \bar{C}_1^* et \bar{C}_2^* les deux matrices ainsi définies, nous avons :

$$\bar{c}_1^* \approx \begin{bmatrix} 0,69 & -0,73 \\ -1,11 & -0,21 \\ -1,51 & 0,46 \\ 0,86 & 1,16 \\ 0,26 & -0,95 \\ 0,82 & 0,27 \end{bmatrix} ; \quad \bar{c}_2^* \approx \begin{bmatrix} 0,88 & 0,99 \\ 0,49 & -1,17 \\ -1,37 & 0,18 \\ \hline 0,90 & 1,61 \\ 0,79 & -0,32 \\ -0,45 & -0,80 \\ -1,58 & 0,64 \\ \hline 0,69 & -0,09 \\ -1,37 & 0,18 \end{bmatrix} .$$

Remarque : Dans la matrice des données X , on peut remarquer que l'on a $X_1^3 = X_1^1 + X_2^1$ et $X_2^3 = X_3^1$; avec les liaisons habituelles entre indicatrices, cela nous donne 4 relations indépendantes entre les 9 variables indicatrices X_i^j , d'où une matrice de rang 5 et l'existence de 5 valeurs propres non triviales. Par ailleurs, au niveau des composantes principales, on peut vérifier que l'on retrouve bien $(\bar{c}_2)_8^k = \frac{1}{2} [(\bar{c}_2)_1^k + (\bar{c}_2)_2^k]$ et $(\bar{c}_2)_9^k = (\bar{c}_2)_3^k$, pour tout k entre 1 et 5.

On peut maintenant construire le graphique présenté à la page suivante.



Les contributions relatives des individus à l'inertie de leur nuage sont :

1 : 12,5% 2 : 16,7% 3 : 25% 4 : 20,8 %
5 : 12,5% 6 : 12,5% .

Pour les modalités, on a :

X_1^1 : 11,1% X_2^1 : 11,1% X_3^1 : 11,1%
 X_1^2 : 13,9% X_2^2 : 11,1% X_3^2 : 11,1% X_4^2 : 13,9%
 X_1^3 : 5,6% X_2^3 : 11,1% .

Les conditions relatives des individus à la détermination des axes sont :

axe 1 : 1 : 8,7% 2 : 22,6% 3 : 41,5%
 4 : 13,6% 5 : 1,2% 6 : 12,4% ;
axe 2 : 1 : 17,2% 2 : 1,3% 3 : 6,9%
 4 : 43,1% 5 : 29,2% 6 : 2,3% .

Les contributions relatives des modalités à la détermination des axes sont :

axe 1 : X_1^1 : 9,5% X_2^1 : 3,0% X_3^1 : 22,9%
 X_1^2 : 5,0% X_2^2 : 7,6% X_3^2 : 2,5% X_4^2 : 15,2%
 X_1^3 : 11,5% X_2^3 : 22,9% ;
axe 2 : X_1^1 : 20,9% X_2^1 : 29,2% X_3^1 : 0,7%
 X_1^2 : 27,6% X_2^2 : 2,2% X_3^2 : 13,8% X_4^2 : 4,4%
 X_1^3 : 0,3% X_2^3 : 0,7% .

6. A.F.C. D'UN TABLEAU DE BURT RELATIF A PLUS DE DEUX VARIABLES QUALITATIVES

6.1. Données et notations.

Les données sont les mêmes qu'au point 5. et nous prenons également les mêmes notations.

Soit B le tableau de BURT $c \times c$ relatif à l'ensemble des p variables considérées : $B = {}^t X X$.

Les effectifs marginaux des lignes et des colonnes de B valent

$p n_1^1, \dots, p n_{c_1}^1, \dots, p n_1^p, \dots, p n_{c_p}^p$, l'effectif total valant np^2 .

$$\tilde{T} = B ; \quad \tilde{D}_1 = \tilde{D}_2 = \frac{1}{p} \begin{bmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_p \end{bmatrix} = \frac{1}{p} \Delta ; \quad \tilde{D}_1^{-1} = \tilde{D}_2^{-1} = p \Delta^{-1} .$$

$$\tilde{A}_1 = \tilde{A}_2 = \frac{1}{np^2} \tilde{D}_1^{-1} B = \frac{1}{np} \Delta^{-1} B = \frac{1}{p} {}^t S \quad (\text{voir le 5.2.}) .$$

6.2. A.C.P. des profils-lignes (ou des profils-colonnes) de B .

Il est clair que, là encore, les deux A.C.P. donneront les mêmes résultats ; on n'envisage donc que l'A.C.P. des profils-lignes.

• Matrice à diagonaliser :

C'est ${}^t \tilde{A}_1 \quad {}^t \tilde{A}_2 = \left(\frac{1}{p} S \right)^2$.

Nous avons diagonalisé ${}^t \tilde{A}_1 \quad {}^t \tilde{A}_2 = \frac{1}{p} S$ en 5.2. ; nous allons donc trouver ici les mêmes vecteurs propres (éventuellement, à une homothétie près) ,

les valeurs propres étant élevées au carré. Le nombre de valeurs propres non nulles de ${}^t \tilde{A}_1 \quad {}^t \tilde{A}_2$, en dehors de la valeur propre triviale 1, est donc ici la même que précédemment, soit m .

• Matrice des vecteurs propres :

On a $\tilde{D}_2^{-1} = p \Delta^{-1} = \bar{D}_2^{-1}$; les vecteurs propres \tilde{D}_2^{-1} -orthonormés de ${}^t \tilde{A}_1 \quad {}^t \tilde{A}_2$ sont donc exactement identiques aux vecteurs propres \bar{D}_2^{-1} -orthonormés de ${}^t \tilde{A}_1 \quad {}^t \tilde{A}_2$; la matrice \tilde{U}_2 , $c \times m$, de ces vecteurs propres disposés en colonnes est donc égale à \bar{U}_2 .

• Matrice des composantes principales :

$$\begin{aligned} \text{Elle s'écrit : } \tilde{C}_1 &= \tilde{A}_1 \tilde{D}_2^{-1} \tilde{U}_2 = {}^t S \Delta^{-1} \bar{U}_2 = \frac{1}{n} \Delta^{-1} B \Delta^{-1} \bar{U}_2 \\ &= \Delta^{-1} S \bar{U}_2 = p \Delta^{-1} \bar{U}_2 N \quad (\text{voir (3) en 5.2.}) \\ &= \bar{C}_2 N^{1/2} \quad (\text{voir 5.3.}) . \end{aligned}$$

$$\text{On a donc finalement : } \tilde{C}_1 = \tilde{C}_2 = \bar{C}_2 N^{1/2} = \begin{bmatrix} C_1 \\ \vdots \\ C_p \end{bmatrix} N^{1/2} ,$$

où $N^{1/2} = \text{diag}(\sqrt{v_1}, \dots, \sqrt{v_m})$.

- Représentation des modalités des variables X^j :

Chaque modalité i de chaque variable X^j est représentée au moyen de la i -ième ligne de $C_j N^{1/2} = p D_j^{-1} U_j N$.

Il ne s'agit donc pas de la même représentation que celle donnée par l'A.F.C. du tableau disjonctif complet, puisque la coordonnée sur le k -ième axe ($1 \leq k \leq m$) est ici multipliée par le facteur $\sqrt{v_m}$; toutefois, les deux représentations seront en général très voisines.

Ce résultat est tout à fait analogue à celui obtenu dans le cas $p=2$.

6.3. Bilan.

- Graphique obtenu :

Comme dans le cas $p=2$, la représentation simultanée est ici inutile puisque $\tilde{C}_1 = \tilde{C}_2$; le seul graphique réalisé à partir de \tilde{C}_1 (ou de \tilde{C}_2) fournit la représentation de l'ensemble des modalités de toutes les variables. On notera bien sûr que l'on n'a pas dans ce cas de représentation des individus. Cette technique constitue encore une fois une généralisation naturelle de celle vue dans le cas $p=2$, laquelle était très voisine de l'A.F.C. classique de la table de contingence T ; elle constitue la *deuxième façon de pratiquer l'Analyse Factorielle des Correspondances Multiple*.

- Autres résultats :

En plus du graphique, on obtient là encore des indicateurs de qualité sans signification concrète et des contributions relatives ayant la même signification que celle indiquée en 4.4.. Les contributions relatives des modalités de chaque variable à la détermination des axes sont encore une fois les mêmes que dans l'A.F.C. du tableau disjonctif complet.

- Pratique de l'Analyse Factorielle des Correspondances Multiples:

Les représentations des modalités des variables sont très voisines dans les deux façons de procéder ; il ne peut donc s'agir d'un critère de choix entre les deux. La première méthode (A.F.C. du tableau disjonctif complet) permet, seule, de représenter les individus et est donc a priori plus intéressante ; mais, dès que l'on a un nombre important d'individus, le graphique est illisible et il est alors préférable d'utiliser la seconde méthode (A.F.C. du tableau de BURT). Les Analyses de Correspondances Multiples étant en général utilisées sur des données nombreuses, c'est le plus souvent par l'A.F.C. du tableau de BURT que l'on réalise ces analyses.

6.4. Exemple.

Reprenons l'exemple traité en 5.5.; le tableau de BURT, avec ses marges, est le suivant :

B :

	x_1^1	x_2^1	x_3^1	\vdots	x_1^2	x_2^2	x_3^2	x_4^2	\vdots	x_1^3	x_2^3	
x_1^1	2	0	0	\vdots	1	1	0	0	\vdots	2	0	6
x_2^1	0	2	0	\vdots	0	1	1	0	\vdots	2	0	6
x_3^1	0	0	2	\vdots	0	0	1	1	\vdots	0	2	6
x_1^2	1	0	0	\vdots	1	0	0	0	\vdots	1	0	3
x_2^2	1	1	0	\vdots	0	2	0	0	\vdots	2	0	6
x_3^2	0	1	1	\vdots	0	0	2	0	\vdots	1	1	6
x_4^2	0	0	1	\vdots	0	0	0	1	\vdots	0	1	3
x_1^3	2	2	0	\vdots	1	2	1	0	\vdots	4	0	12
x_2^3	0	0	2	\vdots	0	0	1	1	\vdots	0	2	6
	6	6	6	\vdots	3	6	6	3	\vdots	12	6	54

La matrice à diagonaliser est le carré de ${}^t\bar{A}_1 {}^t\bar{A}_2$, diagonalisée dans le cas précédent (5.5.) ; elle est 9×9 et s'écrit $\frac{1}{9} (B \Gamma^{-1})^2$, où

$\Gamma = \text{diag}(2, 2, 2, 1, 2, 2, 1, 4, 2)$.

Ses valeurs propres sont :

$$\rho_0 = v_0^2 = 1 \quad \rho_1 = v_1^2 \approx 0,834 \quad \rho_2 = v_2^2 \approx 0,270$$

$$\rho_3 = v_3^2 \approx 0,111 \quad \rho_4 = v_4^2 \approx 0,032 \quad \rho_5 = v_5^2 \approx 0,003$$

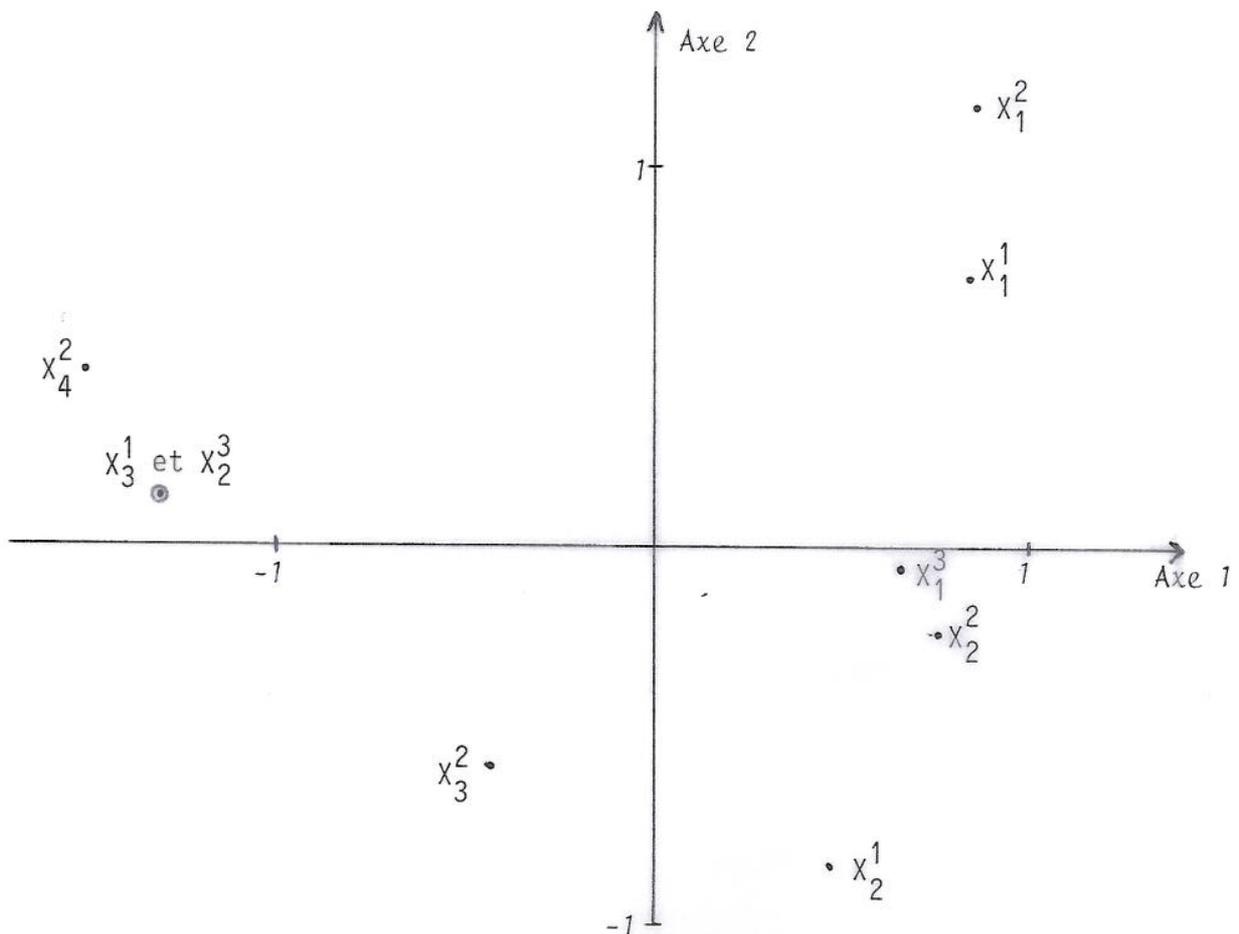
$$\rho_6 = \rho_7 = \rho_8 = 0 .$$

Les matrices des composantes principales s'écrivent :

$$\tilde{C}_1^* = \tilde{C}_2^* \approx \begin{bmatrix} 0,84 & 0,71 \\ 0,47 & -0,84 \\ -1,31 & 0,13 \\ \hline 0,86 & 1,16 \\ 0,76 & -0,23 \\ -0,43 & -0,58 \\ -1,51 & 0,46 \\ \hline 0,66 & -0,06 \\ -1,31 & 0,13 \end{bmatrix}$$

On peut encore remarquer que $(\tilde{C}_1)_8^k = \frac{1}{2} [(\tilde{C}_1)_1^k + (\tilde{C}_1)_2^k]$ et que $(\tilde{C}_1)_9^k = (\tilde{C}_1)_3^k$, pour tout k compris entre 1 et 5.

On déduit de \tilde{C}_1^* le graphique suivant :



On constate qu'il est très proche de celui obtenu en 5.5. . Les contributions relatives de l'ensemble des modalités à l'inertie de leur nuage valent :

$$\begin{array}{llll} X_1^1 : 11,1\% & X_2^1 : 8,9\% & X_3^1 : 15,6\% & \\ X_1^2 : 11,1\% & X_2^2 : 8,9\% & X_3^2 : 7,8\% & X_4^2 : 13,3\% \\ X_1^3 : 7,8\% & X_2^3 : 15,6\% & & \end{array}$$

Les contributions relatives des modalités à la détermination des axes sont les mêmes qu'en 5.5..

7. APPLICATIONS DE L'ANALYSE CANONIQUE A L'A.F.C.

On se propose dans ce paragraphe de considérer l'Analyse Canonique (A.C.) selon deux conceptions, et de l'appliquer dans chaque cas à des variables qualitatives afin de comparer les résultats obtenus, notamment les représentations graphiques, avec ceux de l'A.F.C.. La première conception est l'approche classique de l'A.C., la seconde l'approche envisagée par J.D. CARROL pour généraliser l'A.C. (voir [4]) et reprise par G. SAPORTA dans [9].

Les résultats les plus classiques ne seront pas démontrés ; on pourra se reporter à [3] pour la première approche et à [9] pour la seconde.

7.1. Données, notations et problème général de l'Analyse Canonique.

(X_1^1, \dots, X_ℓ^1) et (X_1^2, \dots, X_c^2) sont deux ensembles de variables quantitatives observées sur n individus notés ω et munis de poids p_ω ($\forall \omega = 1, \dots, n : p_\omega > 0 ; \sum_{\omega=1}^n p_\omega = 1$).

On note X_1 la matrice $n \times \ell$ des observations centrées des variables X_i^1 ($i=1, \dots, \ell$) et X_2 la matrice $n \times c$ des observations centrées des variables X_j^2 ($j=1, \dots, c$).

On cherche à étudier les liaisons linéaires entre les deux ensembles de variables ; pour cela, on considère l'espace vectoriel \mathbb{R}^n muni de la base canonique et de la métrique associée à la matrice $D = \text{diag}(p_1, \dots, p_n)$ sur cette base (métrique des poids).

On peut ensuite voir le problème de deux façons distinctes, mais équivalentes.

7.2. Première conception de l'Analyse Canonique.

• Principe :

Soit W_1 (respectivement W_2) le sous-espace vectoriel de \mathbb{R}^n engendré par les vecteurs associés à X_1^1, \dots, X_ℓ^1 (respectivement X_1^2, \dots, X_c^2) ; ces deux s.e.v. sont supposés de dimensions respectives ℓ et c ; et l'on suppose de plus $\ell \leq c$.

On cherche dans W_1 une base D-orthonormée ξ^1, \dots, ξ^ℓ et dans W_2 une autre base D-orthonormée η^1, \dots, η^c telles que :

- (ξ^1, η^1) soit le couple le plus corrélé de $W_1 \times W_2$;
- (ξ^2, η^2) soit le couple de $W_1 \times W_2$, ξ^2 orthogonal à ξ^1 et η^2 orthogonal à η^1 , le plus corrélé ;
- et ainsi de suite ...

• Solution :

Les (ξ^k, η^k) , $k=1, \dots, \ell$, sont appelées les variables canoniques ; si P_1 (respectivement P_2) est la matrice carrée d'ordre n associée à la projection D-orthogonale sur W_1 (respectivement W_2), on montre alors que :

ξ^k (respectivement η^k) est vecteur propre de $P_1 P_2$ (respectivement $P_2 P_1$) associé à λ_k ($\lambda_1 \geq \dots \geq \lambda_\ell > 0$ sont les valeurs propres non nulles communes à $P_1 P_2$ et $P_2 P_1$).

Remarques :

$$- P_1 = X_1 V_{11}^{-1} {}^t X_1 D \quad \text{et} \quad P_2 = X_2 V_{22}^{-1} {}^t X_2 D ,$$

$$\text{où } V_{11} = {}^t X_1 D X_1 \quad \text{et} \quad V_{22} = {}^t X_2 D X_2 .$$

- $\eta^{\ell+1}, \dots, \eta^c$ sont choisis arbitrairement dans W_2 pour constituer avec η^1, \dots, η^ℓ une base D-orthonormée.

• Propriété :

$$\text{Posons : } A_1 = V_{11}^{-1} V_{12} , \quad \text{où } V_{12} = {}^t X_1 D X_2 ,$$

$$\text{et } A_2 = V_{22}^{-1} V_{21} , \quad \text{où } V_{21} = {}^t X_2 D X_1 = {}^t V_{12} .$$

Dans \mathbb{R}^ℓ muni de la base canonique et de la métrique associée à V_{11}^{-1} sur cette base, soit u_1^k le vecteur propre V_{11}^{-1} -normé de ${}^t A_2 {}^t A_1$ associé à la k-ième valeur propre μ_k ($k=1, \dots, \ell$).

On peut alors vérifier : $\xi^k = X_1 V_{11}^{-1} u_1^k$ et $\mu_k = \lambda_k$.

De la même manière, on vérifie que $\eta^k = X_2 V_{22}^{-1} u_2^k$, où les u_2^k sont les vecteurs propres V_{22}^{-1} -orthonormés de ${}^t A_1 {}^t A_2$, associés aux valeurs propres non nulles λ_k ($k=1, \dots, \ell$).

• Représentation des variables :

On représente les variables X_i^1 au moyen des coordonnées des x_i^1 (vecteurs les représentant dans \mathbb{R}^n) sur la base des ξ^k , c'est-à-dire au moyen des

$$\begin{aligned} \langle x_i^1, \xi^k \rangle_D &= {}^t [x_i^1] D [\xi^k] \\ &= {}^t [x_i^1] D X_1 V_{11}^{-1} u_1^k . \end{aligned}$$

Soit U_1 la matrice $\ell \times \ell$ donnant les u_1^k (V_{11}^{-1} -orthonormés) en colonnes ; les coordonnées des variables X_i^1 par rapport aux ξ^k s'obtiennent dans les lignes de la matrice ${}^t X_1 D X_1 V_{11}^{-1} U_1 = V_{11} V_{11}^{-1} U_1 = U_1$.

Pareillement, on peut représenter sur un autre graphique en dimension ℓ les variables X_j^2 au moyen des lignes de la matrice U_2 ($c \times \ell$) dont la k -ième colonne est u_2^k .

Ensuite, si l'on souhaite (et c'est en général le cas) disposer d'une représentation graphique simultanée des deux ensembles de variables, on peut soit projeter les x_j^2 dans W_1 , soit projeter les x_i^1 dans W_2 . Dans le premier cas, la coordonnée de x_j^2 relativement à ξ^k est :

$$\langle P_1 x_j^2, \xi^k \rangle_D = {}^t_{[x_j^2]} {}^t_{P_1} D X_1 V_{11}^{-1} [u_1^k] ;$$

on utilise donc la matrice :

$$\begin{aligned} {}^t_{X_2} {}^t_{P_1} D X_1 V_{11}^{-1} U_1 &= {}^t_{X_2} D X_1 V_{11}^{-1} {}^t_{X_1} D X_1 V_{11}^{-1} U_1 \\ &= V_{21} V_{11}^{-1} U_1 = {}^t_{A_1} U_1 . \end{aligned}$$

On utilisera donc dans ce premier cas les lignes de U_1 pour représenter les variables X_i^1 et celles de ${}^t_{A_1} U_1$ pour représenter les variables X_j^2 . Dans le second cas, on aurait respectivement recours à ${}^t_{A_2} U_2$ et à U_2 .

Cette conception de l'A.C. présente donc l'inconvénient majeur d'aboutir à deux représentations simultanées distinctes des variables, sans qu'il soit possible d'en privilégier une.

• Représentation des individus :

Il est encore possible en A.C. de faire deux représentations des individus observés. Elles se font l'une dans \mathbb{R}^ℓ et l'autre dans \mathbb{R}^c , ne sont pas identiques et ne peuvent en principe pas se superposer à la représentation des variables. Nous ne les développons pas ici car elles ne seront pas utilisées par la suite, mais on pourra se reporter à [3].

7.3. Application à l'Analyse Factorielle des Correspondances.

• Principe :

On remplace les variables X_i^1 par les indicatrices des modalités de X^1 et les X_j^2 par celles des modalités de X^2 ; les matrices X_1 et X_2 sont alors des tableaux d'indicatrices dont on effectue l'A.C. .

Remarque : Les matrices X_1 et X_2 correspondront ici à des données non centrées (cela était déjà le cas au paragraphe 2. pour les matrices de profils A_1 et A_2) ; la technique s'en trouve simplifiée et l'on peut vérifier que les résultats sont équivalents (à un vecteur constant près).

• Représentation des modalités de X^1 :

Elle se fait au moyen des lignes de U_1 .

On notera ici que :

$$V_{11} = {}^t X_1 D X_1 = \frac{1}{n} {}^t X_1 X_1 = D_1 \quad ; \quad V_{11}^{-1} = D_1^{-1} \quad ;$$

$$V_{22} = D_2 \quad ; \quad V_{22}^{-1} = D_2^{-1} \quad ;$$

$$V_{12} = {}^t X_1 D X_2 = \frac{1}{n} {}^t X_1 X_2 = \frac{1}{n} T \quad ; \quad V_{21} = \frac{1}{n} {}^t T \quad ;$$

$A_1 = \frac{1}{n} D_1^{-1} T$ est le A_1 habituel de l'A.F.C. (tableau des profils-lignes) ;

$A_2 = \frac{1}{n} D_2^{-1} {}^t T$ est le A_2 habituel de l'A.F.C. (tableau des profils-colonnes).

Les vecteurs propres V_{11}^{-1} -orthonormés de ${}^t A_2 {}^t A_1$ sont donc les vecteurs propres D_1^{-1} -orthonormés de la matrice diagonalisée en A.F.C. ; nous les avons notés u_1^k au point 2.4. ; il y a donc équivalence entre les deux notations .

Rappelons qu'en A.F.C. classique, les modalités de X^1 sont représentées au moyen des lignes de C_1 et que nous avons vu (en 2.5.) que $C_1 = D_1^{-1} U_1 L^{1/2}$; les deux représentations sont donc différentes.

• Représentation des modalités de X^2 :

Soit U_2 la matrice $c \times (\ell - 1)$ des vecteurs propres V_{22}^{-1} -orthonormés de ${}^t A_1 {}^t A_2$; il est clair que $V_{22}^{-1} = D_2^{-1}$ et que U_2 est égale à la matrice ainsi notée en A.F.C. ; le résultat est donc semblable à celui obtenu pour les modalités de X^1 .

• Représentations simultanées :

Il est encore possible ici d'en faire 2. Dans la première, les modalités de X^1 seront représentées au moyen des lignes de $U_1 = D_1 C_1 L^{-1/2}$ et celles de X^2 au moyen des lignes de ${}^t A_1 U_1 = U_2 L^{1/2}$ (on a vu en 2.4. que $U_2 = {}^t A_1 U_1 L^{-1/2}$) = $D_2 C_2$. Dans la seconde, on utilisera les lignes respectives de $U_1 L^{1/2} = D_1 C_1$ et de $U_2 = D_2 C_2 L^{-1/2}$. Ces représentations graphiques sont donc toutes deux différentes de la représentation simultanée brute obtenue par l'A.F.C. de la table de contingence T , et réalisée à partir des lignes de C_1 et de C_2 ; en

particulier, on passe de C_1 à U_1 ou de C_2 à U_2 en effectuant à la fois une pré-multiplication et une post-multiplication, ce qui peut changer sensiblement le graphique qui en découle.

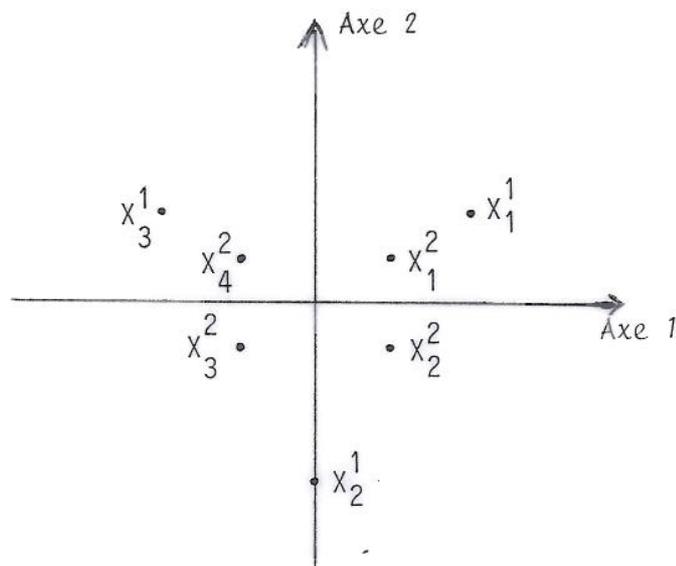
• Exemple :

Nous considérons une nouvelle fois l'exemple introduit en 2.7, et nous réalisons la première représentation simultanée.

On avait obtenu : $U_1 \approx \begin{bmatrix} 0,41 & 0,24 \\ 0 & -0,47 \\ -0,41 & 0,24 \end{bmatrix}$;

il vient par ailleurs : ${}^t_{A_1} U_1 \approx \begin{bmatrix} 0,20 & 0,12 \\ 0,20 & -0,12 \\ -0,20 & -0,12 \\ -0,20 & 0,12 \end{bmatrix}$.

D'où le graphique :



7.4. Deuxième conception de l'Analyse Canonique.

• Principe :

On cherche le vecteur z^1 de \mathbb{R}^n le plus proche à la fois de W_1 et de W_2 ; on prend pour critère $R^2(z^1, W_1) + R^2(z^1, W_2)$ maximum, où R^2 désigne le carré du coefficient de corrélation linéaire multiple, et l'on impose à z^1 d'être D-normé ; on recommence en cherchant z^2 D-normé et D-orthogonal à z^1 , le plus proche de W_1 et de W_2 , et ainsi de suite ...

• Solution :

Les z^k , $k=1, \dots, n$, sont les vecteurs propres D-orthonormés de $P_1 + P_2$ associés aux valeurs propres décroissantes : $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n \geq 0$. (voir [9]).

• Relation entre z^k , ξ^k et η^k ($1 \leq k \leq \ell$) :

On peut montrer que $\xi^k = a P_1 z^k$ et $\eta^k = a P_2 z^k$, a étant une constante réelle. Par ailleurs : $\xi^k + \eta^k = a(P_1 + P_2)z^k = a \alpha_k z^k \Rightarrow z^k = \frac{\xi^k + \eta^k}{a \alpha_k}$; comme on peut vérifier $\alpha_k = 1 \pm \sqrt{\lambda_k}$, le signe étant plus pour les plus grandes valeurs propres et $a = \sqrt{\frac{2}{\alpha_k}}$, il vient : $z^k = [2(1 + \sqrt{\lambda_k})]^{-1/2} (\xi^k + \eta^k)$.

• Représentation des variables :

Elle se fait sur le sous-espace de \mathbb{R}^n engendré par les z^k ; coordonnées de X_i^1 : $\langle x_i^1, z^k \rangle_D$.

En écrivant alors z^k en fonction de ξ^k et η^k , en écrivant $\eta^k = \frac{1}{\sqrt{\lambda_k}} P_2 \xi^k$, en développant P_2 et ξ^k ($X_1 V_{11}^{-1} u_1^k$; voir 7.2), et en remarquant que

${}^t [x_i^1]_D [\xi^k] = (U_1)_i^k$ (d'après 7.2), on peut finalement vérifier que la matrice $\ell \times \ell$ donnant dans ses lignes les coordonnées des variables X_i^1 relativement aux z^k s'écrit :

$$\frac{1}{\sqrt{2}} U_1 (I_\ell + L^{1/2})^{1/2} = U_1 N^{1/2}, \text{ avec } N = \frac{1}{2} \text{diag} (\alpha_1, \dots, \alpha_\ell),$$

définie par analogie avec le paragraphe 3..

De la même manière, la matrice $c \times \ell$ donnant dans ses lignes les coordonnées des variables X_j^2 relativement aux z^k est : $U_2 N^{1/2}$.

• Comparaison avec la première conception :

La présentation ci-dessus de l'Analyse Canonique, équivalente à la précédente puisque les résultats se correspondent, présente l'avantage d'aboutir à une représentation graphique unique et naturelle des deux ensembles de variables, dans

un espace "intermédiaire" entre W_1 et W_2 et ne privilégiant aucun de ces deux espaces. Bien entendu, cette représentation diffère des deux représentations simultanées obtenues en 7.2..

• Représentation des individus :

Cette conception de l'Analyse Canonique présente également l'avantage de permettre la définition d'un unique espace d'individus, $\mathbb{R}^{\ell+c}$, dans lequel on peut faire sans ambiguïté leur représentation graphique. Nous ne développons pas ce point ici, car il sera repris de façon plus générale au paragraphe suivant ; notons simplement qu'on trouve les coordonnées des individus (par rapport à une base appropriée d'un sous-espace vectoriel de $\mathbb{R}^{\ell+c}$, de dimension ℓ) dans les lignes de la matrice :

$$\frac{1}{\sqrt{2}} [X_1 \mid X_2] \left[\begin{array}{c|c} V_{11}^{-1} & 0 \\ \hline 0 & V_{22}^{-1} \end{array} \right] \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} .$$

7.5. Application à l'Analyse Factorielle des Correspondances.

On utilise les mêmes principes qu'en 7.3. ; les matrices notées ci-dessus U_1 , U_2 , L et N sont identiques à celles ainsi notées aux paragraphes 2. et 3.. Par conséquent, la représentation simultanée se fait au moyen des matrices :

$$U_1 N^{1/2} = D_1 C_1 L^{-1/2} N^{1/2} ,$$

$$\text{et } U_2 N^{1/2} = D_2 C_2 L^{-1/2} N^{1/2} .$$

On constate que ces matrices ne diffèrent de celles trouvées dans l'A.F.C. du tableau disjonctif complet (paragraphe 3.4.) que par la pré-multiplication par D_1 et D_2 . Toutefois, cela affectant différemment chaque modalité, le graphique peut s'en trouver sensiblement modifié.

La représentation des individus se fait au moyen des lignes de la matrice $\frac{1}{\sqrt{2}} X \Delta^{-1} U$, où X est le tableau disjonctif complet relatif à X^1 et X^2 ,

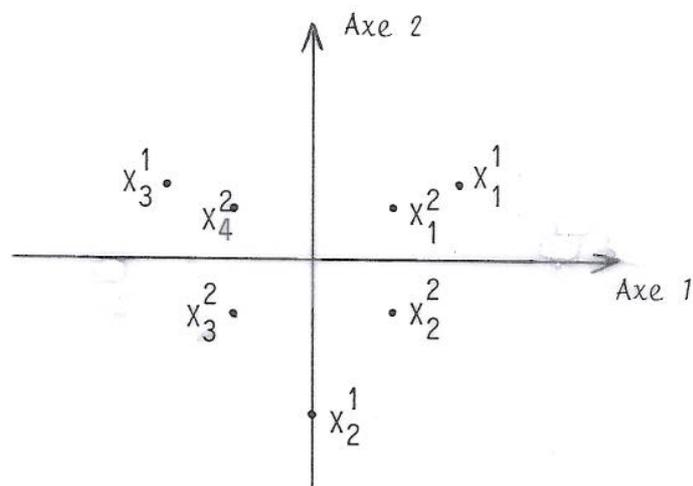
$$\text{où } \Delta^{-1} = \left[\begin{array}{c|c} D_1^{-1} & 0 \\ \hline 0 & D_2^{-1} \end{array} \right] , \quad \text{et où } U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} ,$$

D_1 , D_2 , U_1 et U_2 étant les matrices définies en A.F.C. (voir au paragraphe 2.1. la définition de D_1 et D_2).

• Exemple : Nous reprenons encore l'exemple introduit en 2.7. ; il vient :

$$U_1 N^{1/2} \approx \begin{bmatrix} 0,39 & 0,20 \\ 0 & -0,41 \\ -0,39 & 0,20 \end{bmatrix} \quad \text{et} \quad U_2 N^{1/2} \approx \begin{bmatrix} 0,21 & 0,14 \\ 0,21 & -0,14 \\ -0,21 & -0,14 \\ -0,21 & 0,14 \end{bmatrix} .$$

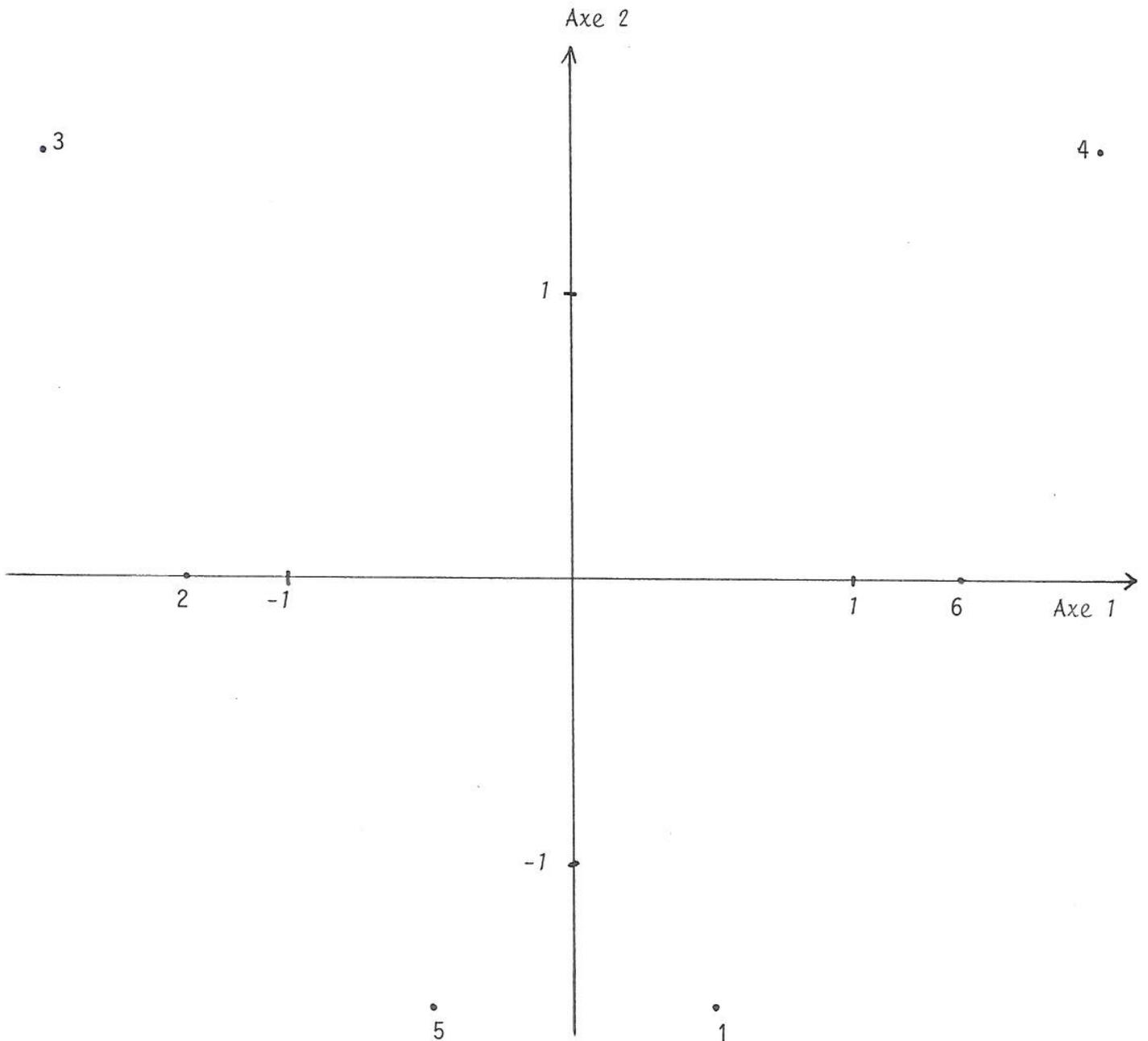
D'où la représentation graphique des variables :



On obtient par ailleurs :

$$\frac{1}{\sqrt{2}} X \Delta^{-1} U \approx \begin{bmatrix} 0,50 & -1,50 \\ -1,37 & 0 \\ -1,87 & 1,50 \\ 1,87 & 1,50 \\ -0,50 & -1,50 \\ 1,37 & 0 \end{bmatrix} .$$

On en déduit la représentation graphique des individus réalisée à la page suivante.



8. APPLICATION DE L'ANALYSE CANONIQUE GENERALISEE (AU SENS DE CARROLL) A
L'ANALYSE FACTORIELLE DES CORRESPONDANCES MULTIPLE

Il existe plusieurs façons de généraliser l'A.C. à plus de deux ensembles de variables quantitatives (voir notamment [9]), toutes se ramenant à l'A.C. classique dans le cas de deux ensembles de variables. Nous ne considérons ici que la généralisation proposée par J.D. CARROLL (voir [4]), car c'est la plus fructueuse au niveau des applications.

8.1. Données, notations et problème général.

On considère p ensembles de variables quantitatives $(X_1^j, \dots, X_{c_j}^j)$, $1 \leq j \leq p$, observées sur n individus ω , $1 \leq \omega \leq n$, affectés de poids p_ω .

On note X_j la matrice $n \times c_j$ des observations centrées du j-ième ensemble, et X la matrice $[X_1 | \dots | X_j | \dots | X_p]$; X est $n \times c$, avec $c = \sum_{j=1}^p c_j$.

En se plaçant alors dans \mathbb{R}^n muni de la métrique des poids, on cherche à étudier les liaisons linéaires entre toutes ces variables.

8.2. Principe de l'Analyse Canonique Généralisée (A.C.G.).

Pour tout j compris entre 1 et p, on note W_j le sous-espace vectoriel de \mathbb{R}^n engendré par les vecteurs associés aux variables X_i^j ($i=1, \dots, c_j$); on supposera que sa dimension est c_j (c'est-à-dire que les X_i^j sont linéairement indépendantes). Notons par ailleurs W la somme de ces sous-espaces (il ne s'agit pas nécessairement d'une somme directe) et m la dimension de W ($m \leq c$).

On cherche alors dans W une base D-orthonormée z^1, \dots, z^m qui soit telle que :

- z^1 est le vecteur D-normé de W maximisant $\sum_{j=1}^p R^2(z, W_j)$ (R désignant toujours le coefficient de corrélation linéaire multiple) ;
- z^2 est le vecteur D-normé de W, D-orthogonal à z^1 , maximisant $\sum_{j=1}^p R^2(z, W_j)$;
- et ainsi de suite ...

On pourra alors faire une représentation graphique *simultanée* de l'ensemble des variables X_i^j sur les sous-espaces de W engendrés par les deux ou trois premiers vecteurs z^k .

8.3. Solutions.

- Les vecteurs z^k ($k=1, \dots, m$) :

Ils sont vecteurs propres de la matrice $\sum_{j=1}^p P_j$ associés aux valeurs propres rangées par ordre décroissant ($\gamma_1 \geq \dots \geq \gamma_m > 0$). P_j est la matrice carrée d'ordre n associée, sur la base canonique de \mathbb{R}^n , au projecteur D-orthogonal sur W_j ; elle vérifie : $P_j = X_j V_{jj}^{-1} {}^t X_j D$, où V_{jj} est la matrice des variances-covariances des variables X_i^j ($i=1, \dots, c_j$).

Il s'ensuit : $\sum_{j=1}^p P_j = \left(\sum_{j=1}^p X_j V_{jj}^{-1} {}^t X_j \right) D = X V^{-1} {}^t X D$,

où
$$V = \begin{bmatrix} V_{11} & & & 0 \\ & \ddots & & \\ & & V_{22} & \\ 0 & & & \ddots \\ & & & & V_{pp} \end{bmatrix} .$$

La matrice V est carrée d'ordre c , symétrique et définie-positive, chaque matrice V_{jj} l'étant puisque les variables X_i^j , $i=1, \dots, c_j$, sont supposées linéairement indépendantes ; V est donc inversible, V^{-1} étant également symétrique et définie-positive.

• Propriété :

Soit $H = {}^t X D X$; H est constituée de p^2 blocs notés V_{gj} ($1 \leq g \leq p$, $1 \leq j \leq p$) et tels que $V_{gj} = {}^t X_{g_j} D X_j$; H est donc symétrique ; d'après les hypothèses faites, son rang est m .

Dans \mathbb{R}^C muni de la base canonique et de la métrique associée à la matrice V^{-1} , on appelle $u^1, \dots, u^k, \dots, u^m$ les vecteurs propres V^{-1} -orthonormés de HV^{-1} associés aux valeurs propres rangées par ordre décroissant ($\beta_1 \geq \dots \geq \beta_m > 0$) ; la V^{-1} -symétrie de HV^{-1} assure l'existence de c valeurs propres réelles non négatives, m seulement étant non nulles puisque HV^{-1} est aussi de rang m .

On a alors la propriété suivante :

$$\forall k = 1, \dots, m : \begin{cases} \beta_k = \gamma_k \\ z^k = \frac{1}{\sqrt{\gamma_k}} X V^{-1} [u^k] \end{cases} .$$

En effet, $(\sum_{j=1}^p p_j) X V^{-1} [u^k] = X V^{-1} H V^{-1} [u^k] = \beta_k X V^{-1} [u^k]$, ce qui établit

l'égalité deux à deux des valeurs propres β_k et γ_k , ainsi que la colinéarité de z^k et de $X V^{-1} [u^k]$; comme on vérifie simplement $\|X V^{-1} [u^k]\|_D^2 = \gamma_k$, le résultat est établi.

• Représentation des variables :

Dans W , on représente chaque variable X_i^j en utilisant les coordonnées sur la base des z^k du vecteur de \mathbb{R}^n associé à X_i^j ; la coordonnée de X_i^j sur z^k est :

$$\langle X_i^j, z^k \rangle_D = {}^t [X_i^j] D [z^k] = \frac{1}{\sqrt{\gamma_k}} {}^t [X_i^j] D X V^{-1} [u^k] .$$

Notons U^* la matrice $c \times m$ donnant dans sa k -ième colonne les coordonnées de u^k sur la base canonique de \mathbb{R}^C et G la matrice $\text{diag}(\gamma_1, \dots, \gamma_m)$.

Les coordonnées des X_i^j sur les z^k sont donc dans les lignes de la matrice :

$${}^t X D X V^{-1} U^* G^{-1/2} = H V^{-1} U^* G^{-1/2} = U^* G^{1/2} .$$

• Représentation des individus :

L'espace des individus est ici \mathbb{R}^C , chaque individu ω y étant représenté par le vecteur x_ω de coordonnées sur la base canonique :

$$[x_\omega] = {}^t [X_1^1(\omega) - \bar{x}_1^1, \dots, X_{C_p}^p(\omega) - \bar{x}_{C_p}^p] .$$

La représentation des individus se fait donc au moyen des coordonnées des x_ω sur les u^k ($1 \leq \omega \leq n$, $1 \leq k \leq m$) dont l'expression est :

$$\langle x_\omega, u^k \rangle_{V^{-1}} = {}^t [x_\omega] V^{-1} [u^k] .$$

Ce sont donc les lignes de la matrice $XV^{-1}U^*$ qui fournissent ces coordonnées.

Remarques :

1) La matrice $XV^{-1}U^*$ est analogue à la matrice C des composantes principales en A.C.P. qui s'écrit $C = XMU$, où X est la matrice des données centrées, M la matrice définissant la métrique dans l'espace des individus et U la matrice des vecteurs propres M-orthonormés dans l'espace des individus (voir 1.3.).

On peut ainsi vérifier (voir [9]) que dans le cas $c_1 = c_2 = \dots = c_p = 1$, l'A.C.G. est équivalente à l'A.C.P..

2) Il n'y a aucune raison de faire ici une représentation simultanée des individus et des variables.

3) On peut ici, comme dans une A.C.P., mesurer la qualité globale des représentations graphique en dimension s ($1 \leq s \leq m$) par le rapport des s plus grandes valeurs propres à la trace de HV^{-1} (qui vaut c). Toutefois, les données étant ici plus particulières, ce rapport n'est pas aussi significatif ; en particulier, dans le cas de deux variables qualitatives, on retrouve les mêmes valeurs propres que dans l'A.F.C. du tableau disjonctif complet, et l'on sait que ces valeurs propres n'ont pas de signification concrète (voir ci-dessous).

8.4. Application de l'A.C.G. à l'A.F.C. multiple.

• Principe :

On dispose de p variables qualitatives $X^1, \dots, X^j, \dots, X^p$ (X^j ayant c_j modalités) observées sur n individus ω ($\omega = 1, \dots, n$) tous affectés du même poids $\frac{1}{n}$.

Pour tout j ($1 \leq j \leq p$), on définit les c_j indicatrices des modalités de X^j que l'on note X_i^j ($1 \leq i \leq c_j$) et on leur associe X_j , tableau $n \times c_j$ des indicatrices non cen-

trées (là encore, les données qualitatives ne seront pas centrées, les résultats étant équivalents).

On note enfin X le tableau disjonctif complet $[X_1 | \dots | X_j | \dots | X_p]$ de dimension $n \times c$ ($c = \sum_{j=1}^p c_j$).

On applique alors l'A.C.G. à l'ensemble de toutes les indicatrices X_i^j .

• Expression des différentes matrices :

$$V_{jj} = {}^t X_j D X_j = \frac{1}{n} {}^t X_j X_j = \frac{1}{n} \text{diag}(n_{1.}^j, \dots, n_{c_j.}^j) = D_j$$

(matrice ainsi notée en 5.1.) .

$$V = \begin{bmatrix} V_{11} & & 0 \\ & \ddots & \\ 0 & & V_{pp} \end{bmatrix} = \begin{bmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_p \end{bmatrix} = \Delta \text{ (cf. 5.1.) ; } \quad \text{d'où } V^{-1} = \Delta^{-1} .$$

$V_{gj} = {}^t X_g D X_j = \frac{1}{n} {}^t X_g X_j = \frac{1}{n} T_{gj}$, où T_{gj} est la table de contingence relative à X^g disposée en lignes et à X^j disposée en colonnes.

$H = {}^t X D X = \frac{1}{n} {}^t X X = \frac{1}{n} B$, où B est le tableau de BURT des observations.

$H V^{-1} = \frac{1}{n} B \Delta^{-1} = S$, matrice ainsi notée en 5.2..

En reprenant les mêmes notations qu'en 5.2., il vient de façon évidente :

$$U^* = \sqrt{p} \bar{U}_2 \quad \text{et} \quad G = p N .$$

• Représentation des modalités des variables X^j :

$U^* G^{1/2} = p \bar{U}_2 N^{1/2}$; c'est donc les lignes de cette matrice qui serviront à représenter les modalités des variables X^j (X_1^j). Rappelons que c'est la matrice $p \Delta^{-1} \bar{U}_2 N^{1/2}$ qui était utilisée dans l'A.F.C. du tableau disjonctif complet ; il y a donc là encore une différence.

• Représentation des individus :

$$X V^{-1} U^* = \sqrt{p} X \Delta^{-1} \bar{U}_2 = \sqrt{p} \sum_{j=1}^p X_j D_j^{-1} U_j ; \text{ en coefficient } \sqrt{p} \text{ près, on re-}$$

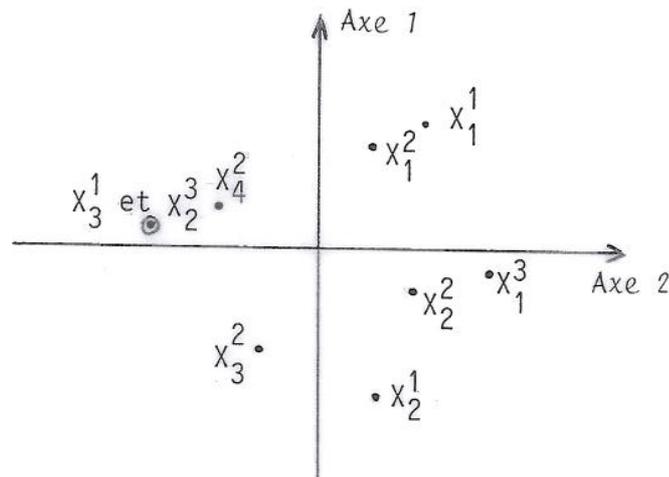
trouve donc ici la même représentation des individus que dans l'A.F.C. du tableau disjonctif complet ; toutefois, dans ce cas, la représentation d'un individu n'a plus rien à voir avec le vecteur moyen associé aux vecteurs représentant les modalités présentées par cet individu ; la représentation superposée des individus et des modalités n'aurait donc ici aucun sens.

• Exemple :

Nous considérons à nouveau ici l'exemple introduit en 5.5. ; on a :
 $U^* G^{1/2} = p \bar{U}_2 N^{1/2} = \Delta \bar{C}_2$; les deux premières colonnes de cette matrice s'écrivent :

0,29	0,33
0,16	-0,39
-0,46	0,06
0,15	0,27
0,26	-0,11
-0,15	-0,27
-0,26	0,11
0,46	-0,06
-0,46	0,06

D'où la représentation graphique des modalités des variables :



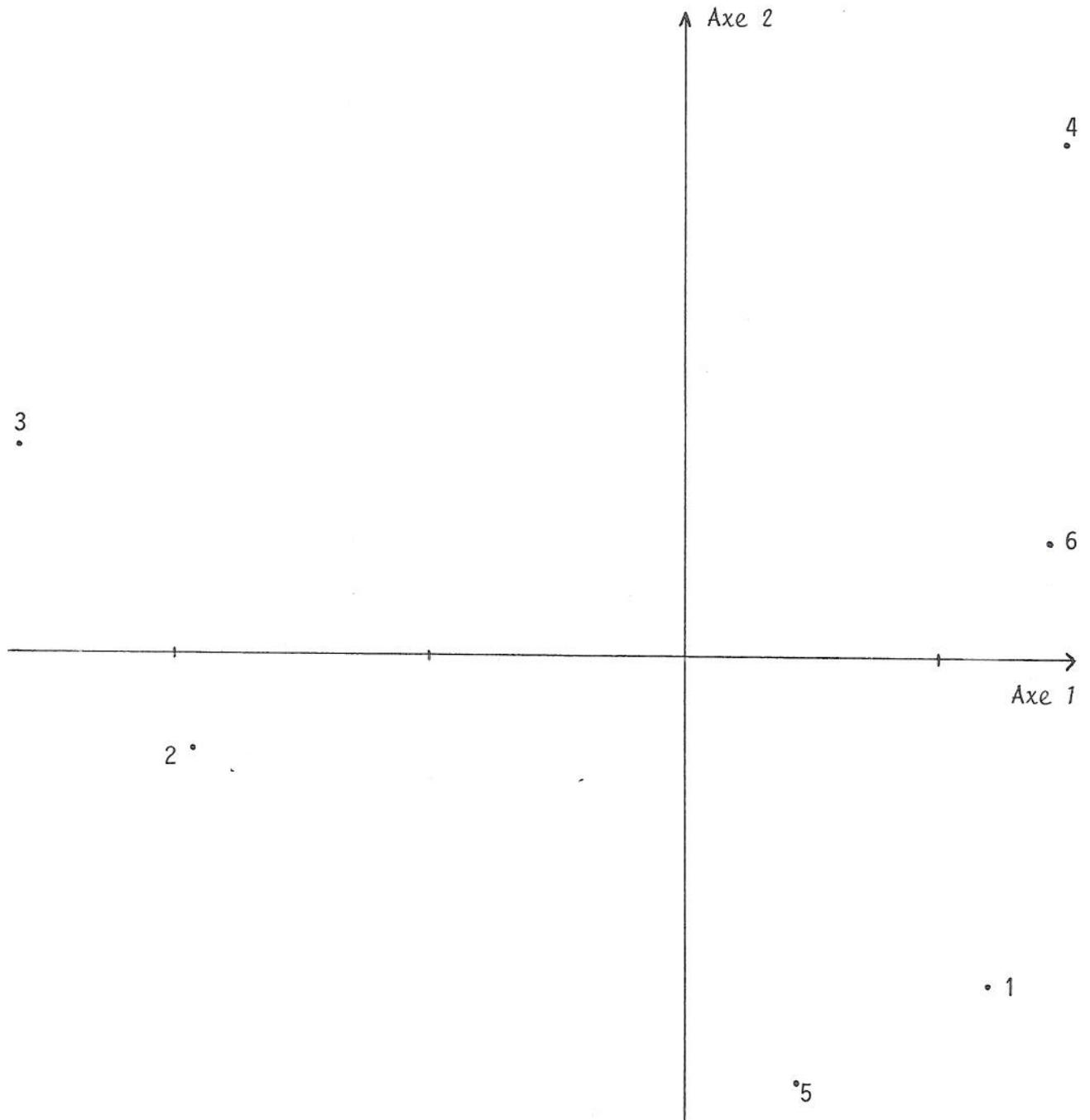
On remarque ici que X_2^3 est toujours confondu avec X_3^1 , alors que X_1^3 n'est plus la moyenne vectorielle de X_1^1 et X_2^1 , mais leur somme vectorielle.

Ce résultat, dû à la prè-multiplication par Δ , ne nous paraît pas satisfaisant ; sur ce point précis, nous trouvons préférable l'A.F.C.M. à l'A.C.G..

D'autre part, $\sqrt{p} \times \Delta^{-1} \bar{U}_2 = \sqrt{p} \bar{C}_1 = \sqrt{3} \bar{C}_1$; les deux premières colonnes s'écrivent :

$$\begin{bmatrix} 1,19 & -1,27 \\ -1,93 & -0,36 \\ -2,61 & 0,81 \\ 1,50 & 2,01 \\ 0,44 & -1,65 \\ 1,43 & 0,46 \end{bmatrix}$$

D'où la représentation graphique des individus :



9. BILAN

Les tableaux qui suivent présentent, pour chacune des méthodes envisagées ici, les résultats obtenus concernant les matrices utilisées pour les représentations graphiques des modalités des variables intervenant ainsi que pour celles des individus, lorsque ces dernières sont possibles ; les notations sont rappelées en-dessous de chaque tableau.

Le premier concerne le cas de deux variables qualitatives et le second le cas général, le premier étant plus complet puisqu'il est possible d'utiliser dans le cas de deux variables des techniques ne se généralisant pas (A.F.C. d'une table de contingence, A.C. classique).

Se pose alors le problème de choisir l'une de ces méthodes au moment de traiter des données. Dans le cas de deux variables, l'A.F.C. a suffisamment fait ses preuves pour que son choix soit tout indiqué. Dans le cas de plus de deux variables qualitatives, le choix est beaucoup plus délicat ; sur le plan théorique, l'A.C.G. semble la méthode la plus fondée, notamment parce que la représentation simultanée des modalités de toutes les variables y apparaît très naturelle ; l'A.F.C. du tableau disjonctif complet est également intéressante puisqu'elle permet une représentation de chaque individu proche du barycentre des modalités qu'il a présentées ; enfin, l'A.F.C. du tableau de BURT, seule méthode ne permettant pas la représentation des individus, fournit une représentation des modalités voisine de la précédente et tout autant justifiée ; c'est essentiellement pour des raisons d'habitude qu'elle est la plus utilisée dans la pratique.

Toutefois, le fait que la représentation des individus soit en général peu importante dans ces méthodes (surtout lorsqu'ils sont très nombreux), ainsi que diverses considérations pratiques (c'est la méthode donnant les résultats les plus proches de l'A.F.C. classique dans le cas $p = 2$; voir la remarque faite dans l'exemple du 8.4. ; ...) nous conduisent à penser que l'A.F.C. sur tableau de BURT est, le plus souvent, *la méthode la plus intéressante* dans la pratique .

9.1. Cas de deux variables qualitatives.

	A.F.C. de la table de contingence ; représentations brutes.	A.F.C. de la table de contingence ; représentations barycentriques.	A.F.C. du tableau disjointif complet	A.F.C. du tableau BURT	A. C. classique (*)	A. C. généralisée
Modalités $\nearrow X_j^1$ $\searrow X_j^2$	$D_1^{-1} U_1 L^{1/2}$ $D_2^{-1} U_2 L^{1/2}$	$D_1^{-1} U_1 L$ $D_2^{-1} U_2 L$	$D_1^{-1} U_1 N^{1/2}$ $D_2^{-1} U_2 N^{1/2}$	$D_1^{-1} U_1 N$ $D_2^{-1} U_2 N$	U_1 U_2	$U_1 N^{1/2}$ $U_2 N^{1/2}$
Individus	—	—	$\frac{1}{2} \times \Delta^{-1} U$	—	—	$\frac{1}{\sqrt{2}} \times \Delta^{-1} U$

A.F.C. signifie Analyse Factorielle des Correspondances et A.C. Analyse Canonique.

T étant la table de contingence initiale, $\ell \times c$, D_1 (resp. D_2) la matrice $\ell \times \ell$ (resp. $c \times c$), diagonale, des fréquences marginales de X^1 (resp. de X^2), on a posé : $A_1 = \frac{1}{n} D_1^{-1} T$ et $A_2 = \frac{1}{n} D_2^{-1} T$ (matrices des profils).

U_1 (resp. U_2) est la matrice $\ell \times (\ell-1)$ (resp. $c \times (c-1)$) des vecteurs propres D_1^{-1} -orthonormés (resp. D_2^{-1} -orthonormés) de ${}^t A_1$ (resp. de ${}^t A_2$) disposés en colonnes, les valeurs propres non nulles correspondantes, $\lambda_1, \dots, \lambda_{\ell-1}$, étant rangées par ordre décroissant.

$$L = \text{diag}(\lambda_1, \dots, \lambda_{\ell-1}) ; N = \frac{1}{2} (I_{\ell-1} + L^{1/2}).$$

X est le tableau disjointif complet, $n \times (\ell+c)$, relatif à X^1 et X^2 .

$$\Delta = \left[\begin{array}{c|c} D_1 & 0 \\ \hline 0 & D_2 \end{array} \right] ; U = \left[\begin{array}{c} U_1 \\ U_2 \end{array} \right].$$

(*) pour une représentation simultanée, on considèrera soit U_1 et $U_2 L^{1/2}$, soit $U_1 L^{1/2}$ et U_2 .

9.2. Cas de plus de deux variables qualitatives.

	A.F.C. du tableau disjonctif complet	A.F.C. du tableau de BURT	A.C. généralisée
Modalités	$\Delta^{-1} U^* G^{1/2}$	$\frac{1}{\sqrt{p}} \Delta^{-1} U^* G$	$U^* G^{1/2}$
Individus	$\frac{1}{\sqrt{p}} X \Delta^{-1} U^*$	_____	$X \Delta^{-1} U^*$

Chaque variable X_i^j ($1 \leq j \leq p$) possède c_j modalités notées X_i^j ($1 \leq i \leq c_j$) ; $c = \sum_{j=1}^p c_j$.

X est le tableau disjonctif complet correspondant, $n \times c$, B le tableau de BURT, D_j la matrice $c_j \times c_j$, diagonale, des fréquences marginales relatives à X_i^j ; on a posé : $\Delta = \begin{bmatrix} D_1 & & 0 \\ & \dots & \\ 0 & & D_p \end{bmatrix}$.

U^* est la matrice $c \times m$ des vecteurs propres Δ^{-1} -orthonormés de $S = \frac{1}{n} B \Delta^{-1}$, disposés en colonnes et associés aux valeurs propres non nulles rangées par ordre décroissant, $\gamma_1 \geq \dots \geq \gamma_m > 0$; $G = \text{diag}(\gamma_1, \dots, \gamma_m)$.

Remarques : 1) Les matrices de vecteurs propres et de valeurs propres définies aux paragraphes 5. et 6. se déduisent des matrices ci-dessus au moyen des relations :

$$\bar{U}_2 = \frac{1}{\sqrt{p}} U^* ; \quad N = \frac{1}{p} G .$$

2) En faisant $p=2$ dans les différents cas ci-dessus, on retrouve les résultats du 9.1. en remarquant

$$U = 2\bar{U}_2 = \sqrt{2} U^* \quad \text{et} \quad N = \frac{1}{2} G .$$

BIBLIOGRAPHIE

- [1] T.W. ANDERSON (1958) : "An introduction to multivariate statistical analysis", Wiley.
- [2] P. BERTIER et J.M. BOUROCHE (1975) : "Analyse des données multidimensionnelles", Presses Universitaires de France.
- [3] F. CAILLIEZ et J.P. PAGES (1976) : "Introduction à l'analyse des données", Smash.
- [4] J.D. CARROLL (1968) : "A generalisation of canonical correlation analysis to three or more sets of variables", Proc. 76th. Conv. Amer. Psych. Assoc., p.227-228.
- [5] H. CAUSSINUS (1983) : "Modèles pour l'analyse en composantes principales", Publications du Laboratoire de Statistique et Probabilités, Université Paul Sabatier, Toulouse.
- [6] P. CAZES, A. BAUMERDER, S. BONNEFOUS et J.P. PAGES (1977) : "Codage et analyse des tableaux logiques ; introduction à la pratique des variables qualitatives", Cahiers du B.U.R.O., n°27.
- [7] R. CEHESSAT (1976) : "Exercices commentés de statistique et informatique appliquées", Dunod.
- [8] L. LEBART, A. MORINEAU, N. TABARD (1977) : "Techniques de la description statistique", Dunod.
- [9] G. SAPORTA (1975) : "Liaison entre plusieurs ensembles de variables et codage de données qualitatives", Thèse de 3ème cycle, Université de Paris VI.

*Tirage : Imprimerie du Laboratoire
de Statistique et Probabilités
Université Paul Sabatier - Toulouse*

Dépot légal : 4ème trimestre 1984, N° 02-84