

55376
1990
19

55376
1990
19

N° d'ordre : 575

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE FLANDRES ARTOIS

pour obtenir

LE TITRE DE DOCTEUR DE L'UNIVERSITÉ

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

par

François LECHEVALIER



L'ANALYSE FACTORIELLE CONJOINTE D'UNE FAMILLE DE TRIPLETS INDEXÉS

Soutenue le 19 Septembre 1990 devant la Commission d'Examen :

Président : G. LE CALVÉ, Université de RENNES II

Rapporteurs : G.SAPORTA, C.N.A.M. PARIS

G. LE CALVÉ, Université de RENNES II

Examineurs : N. VANEECLOO, Université de LILLE I

P. OGER, E.D.F. - E.N.S.A.E. PARIS

C. LANGRAND, Université de LILLE I

B.U. LILLE I



D 030 083928 0

Je tiens à exprimer ma reconnaissance à tous ceux qui m'ont permis de réaliser cette thèse, en particulier à :

Monsieur le Professeur Georges LE CALVÉ qui me fait l'honneur de présider ce jury et a accepté d'en être le rapporteur. Les pertinentes remarques qu'il m'a formulées m'ont permis d'améliorer ce travail sur plusieurs points.

Monsieur le Professeur Gilbert SAPORTA qui a bien voulu être rapporteur de cette thèse, ses conseils éclairés ont contribué à son achèvement.

Monsieur le Professeur Nicolas VANEECLOO, Doyen de la Faculté des Sciences Economiques et Sociales où j'occupe un poste d'assistant, qui, en étant membre du jury, témoigne de l'ouverture grandissante des praticiens ou théoriciens de l'économie aux méthodes Mathématiques et Statistiques.

Monsieur Philippe OGER, Directeur du GRETS à EDF et Professeur à l'ENSAE, qui me fait l'amitié de s'intéresser à cette recherche et de participer au jury.

Monsieur le Professeur Claude LANGRAND qui m'a fait connaître et aimer l'analyse des données. Malgré sa lourde charge de Directeur de l'UFR de Mathématiques, il a su trouver une disponibilité et une patience permanentes pour aider et diriger cette recherche qui doit beaucoup à ses conseils avisés et chaleureux.

Les membres du groupe "données évolutives" de l'ex GRECO 59 du CNRS dont les critiques constructives m'ont stimulé à chacune de nos nombreuses rencontres.

Monsieur François DANJOU qui a partagé beaucoup de mes tourments de chercheur et a su les transformer en encouragements.

Mesdames Arlette LENGAINNE et Véronique TESTELIN qui ont effectué avec compétence l'essentiel de la dactylographie de cette thèse.

Messieurs Carlos SACRÉ et André D'HOINE qui m'ont initié à l'usage du traitement de texte.

Toutes les personnes des services de reprographie des UFR de Mathématiques et d'Economie.

PLAN DETAILLE

	Pages
INTRODUCTION : Les objectifs de la méthode.	5
a) Rappel des propriétés de l'ACP.	
b) Objectifs pour l'infrastructure.	
I - LES METHODES D'INFRASTRUCTURE.	
1.1) Rappel des principales méthodes d'infrastructure.	9
a) Les méthodes pour données non-cubiques.	
b) Les méthodes pour données cubiques.	
1.2) Les infrastructures inexactes.	11
a) L'AFJ1 (ou l'AFM).	
b) L'AFJ2 (ou Statis).	
1.3) Les infrastructures exactes.	13
a) La "superposition".	
b) L'Analyse Procuste.	
1.4) L'AFJ et les autres méthodes.	14
a) Les méthodes d'infrastructure.	
b) Les méthodes de reconstitution des données.	
II - LES TABLEAUX A COMPOSANTES PROPORTIONNELLES.	
2.1) Le Repère Equisecteur.	17
a) Les systèmes équisecteurs.	
b) Signification statistique pour les données cubiques.	
c) Le Repère Equisecteur pour les données non-cubiques.	
2.2) Introduction à l'AFJ.	23
a) Les tableaux proportionnels.	
b) Les tableaux à composantes principales proportionnelles.	
c) Le Repère Déformant.	

III - L'AFJ D'UNE FAMILLE DE TABLEAUX QUELCONQUES.

- 3.1) Caractérisation des tableaux à composantes proportionnelles. 31
- a) Propriétés du Repère Compromis.
 - b) Propriétés du Repère Déformant.
- 3.2) L'Analyse Factorielle conJointe (AFJ). 35
- a) Les composantes conjointes.
 - b) Nombre minimal admissible de composantes.
 - c) L'algorithme de l'AFJ.
- 3.3) Les résultats de l'AFJ. 41
- a) Interprétation de la méthode.
 - b) Applications.
 - c) L'Analyse Procuste.

IV - LES INTRASTRUCTURES SUR TROIS EXEMPLES.

9 figures et 5 tableaux hors texte.

- 4.1) Les accidents. 45
- a) Des données cubiques.(Tab.1)
 - b) L'Intrastructure de l'ACP du Tableau Conjoint Empilé.(Fig.1)
 - c) L'Intrastructure Equisecteur : des composantes principales proportionnelles.(Tab.2, Fig.2)
 - d) L'Intrastructure Compromis et l'effet PLM.(Fig.3 et 4)
- 4.2) Les cantons du Languedoc Roussillon. 58
- a) Les données en effectifs bruts et profils.(Tab.3)
 - b) L'Intrastructure de l'ACP du Tableau Conjoint Empilé.(Fig.5)
 - c) L'Intrastructure Equisecteur et les calculs de l'AFJ.(Tab.4, Fig.6)
 - d) L'Intrastructure Compromis : quasi identique à l'I.D.P.(Fig.7)
 - e) Les cantons avec les effectifs bruts.(Tab.5, Fig.8 et 9)

V - LES INTERSTRUCTURES.

- 5.1) Les interstructures des A_k et des A_k^2 . 73
- 5.2) Les interstructures de tableaux particuliers. 74
- a) Les tableaux proportionnels.
 - b) Les tableaux à composantes principales proportionnelles.
 - c) Les tableaux à composantes proportionnelles.

5.3) Les interstructures dans le cas général.	76
a) Interprétation de INDSCAL dans \mathbf{R}^{n^2} .	
b) L'interstructure de l'AFM.	
c) L'interstructure de Statis.	
CONCLUSION.	85
a) La pratique de l'Analyse Factorielle conJointe.	
b) L'AFJ et l'analyse de données.	
ANNEXES.	89
I) Dualité, Biplot en ACP, application à l'AFC.	A1
II) L'Intrastructure Discriminante.	A6
III) Analyse Procuste et AFJ.	A13
IV) Comparaison des sous-espaces Compromis et Déformant.	A15
ABREVIATIONS ET NOTATIONS.	109
BIBLIOGRAPHIE.	113

0 - INTRODUCTION

Les objectifs de la méthode

Soit une famille de triplets indexés : (X_k, M_k, D) , $k \in K$. Dans chaque triplet, le tableau X_k dimensionné $(n \times p_k)$ contient les observations d'un paquet J_k de variables quantitatives sur un ensemble I d'individus, et q_k est le rang de X_k . Les entiers n , p_k , m désignent le cardinal respectivement de I , J_k , K ; et on note indifféremment $k \in K$ ou $k \in [m]$. On parle de données cubiques quand tous les paquets J_k sont identiques ($J_k = J$) et de données non cubiques dans le cas contraire que nous étudions plus particulièrement. Nous cherchons un positionnement simultané des nuages de points N_k définis dans les espaces Euclidiens $E_k = (\mathbf{R}_{(e_j)}^{p_k}, M_k)$, par les colonnes des tableaux X'_k (transposés de X_k) supposés centrés pour la métrique des poids D . Ce positionnement simultané – nommé **intrastructure** – peut contenir certains transformés des nuages N_k .

L'intrastructure sert à représenter simultanément les nuages dans des sous-espaces de petite dimension, exactement comme c'est le cas pour un seul nuage dans l'ACP d'un tableau. Il paraît donc judicieux, au plan méthodologique, de définir les critères que doit satisfaire une intrastructure, en s'inspirant des résultats de l'ACP.

Nous rappelons ceux-ci, puis nous précisons les deux critères exigibles pour une intrastructure.

a) Rappel des propriétés de l'ACP.

Dans l'ACP d'un triplet (X, M, D) on construit deux triplets équivalents au premier : (C, I, D) et (Z, M, I) .

- Le nuage des individus $\overline{[X']}$, défini par les colonnes du tableau $X'(p \times n)$, est positionné dans l'espace Euclidien $E = (\mathbf{R}_{(e_j)}^p, M)$. Avec (C, I, D) on définit, par exemple dans Cailliez et Pagès (1976), une image euclidienne $(\overline{[C]}, \mathbf{R}_{(u_j)}^q, I)$ équivalente à l'image euclidienne $(\overline{[X]}, \mathbf{R}_{(e_j)}^p, M)$. Si la décomposition en valeur singulière (s.v.d.) de X s'écrit $X = \Gamma \Delta U'$, on a : $C = \Gamma \Delta$ de dimension $(n \times q)$ avec $q = \text{rang } X$, et $Z = \Delta U'$.

• Avec (Z, M, I) , on obtient pour le nuage des variables, une image euclidienne $(\boxed{Z}, \mathbf{R}_{(\gamma_j)}^q, I)$ équivalente à $(\boxed{X}, \mathbf{R}_{(f_i)}^n, D)$. Ce résultat permet de positionner les variables de façon comparable à celle retenue pour les individus. Il permet aussi de définir très simplement le biplot : sur les axes de l'espace $E^q = (\mathbf{R}_{(u_j)}^q, I)$, les coordonnées du vecteur e_j portant un axe initial (et donnant la direction de la variable initiale j) forment la colonne j de : $U'M = \Delta^{-1}ZM$ (voir Annexe I).

On note X_i la ligne i de X et X^j sa colonne j . Alors $x_i = (X_i)'$ définit l'individu i comme vecteur colonne de E (indice de x en bas), et $x^j = X^j$ définit la variable j comme vecteur colonne de F (indice de x en haut) ; de même pour $u_j = U^j$ et $c^j = C^j$. L'indice k des situations échappe à ces règles et figure indifféremment en haut ou en bas. Il ne repère pas une rangée mais spécifie que l'objet qu'il indice est pris en situation k . $\|X\|_{\varphi_{DM}}$ désigne la norme de matrice induite par le produit scalaire $\langle X_1, X_2 \rangle_{\varphi_{DM}} = \text{trace}(X_1' D X_2 M)$ (Glaçon 1981).

Si on complète le théorème de décomposition, par les théorèmes d'approximation (Eckart-Young 1936, Kristof 1970, Le Calvé 1976) on retrouve la propriété d'optimalité suivante de l'ACP :

Le tableau $(n \times s)$ des s premières composantes principales $(c^j)_{j \in [s]}$, noté C_s , définit la projection du nuage $\boxed{C'}$ sur un sous-espace de dimension s , ayant la plus grande inertie. De manière équivalente $\|C - (C_s \ ; \ 0)\|_{\varphi_{DI}}^2$ est minimale pour le tableau C_s . En terme de reconstitution des données : $\|X - B\|_{\varphi_{DM}}^2$, avec $B(n \times p)$ de rang s , est minimale si $B = \sum_{j=1}^s c^j (u_j)'$.

Pour résumer, l'ACP présente les propriétés suivantes :

- obtention d'une image euclidienne équivalente avec (C, I, D) .
- le tableau C_s est le tableau de rang s qui restitue la plus grande part de l'inertie du nuage $\boxed{C'}$.
- la description des variables initiales est optimale au sens de l'inertie d'un nuage non centré (pour une métrique M diagonale).
- le biplot permet d'indiquer la direction des variables initiales dans les plans principaux retenus pour représenter le nuage des individus, et d'orienter ces plans et les axes principaux.

b) Objectifs pour l'intrastructure.

Pour une représentation simultanée de plusieurs nuages il est évidemment intéressant de retrouver ces 4 propriétés et, dans un premier temps, au moins

les deux premières, en remarquant que la seconde n'a de signification que si la première est vérifiée.

En se limitant aux méthodes d'étude des données non-cubiques on constate que cette première propriété n'est pas vérifiée par Statis ou l'AFM, tandis qu'elle peut l'être dans l'Analyse Procuste ou la méthode dite de "superposition des axes principaux d'inertie", ce qui sera montré ultérieurement.

Ces deux dernières méthodes peuvent fournir deux infrastructures totalement différentes. Cela impose un deuxième objectif pour la recherche d'une infrastructure qui est de justifier le positionnement simultané réalisé.

Nous demandons donc à l'infrastructure d'une famille de tableaux indexés de satisfaire deux conditions : être exacte et être significative.

- Elle est **exacte** si pour tout $k \in K$, le triplet (Y_k, I, D) représenté dans l'infrastructure est équivalent au triplet initial (X_k, M_k, D) .
- Elle est **significative** si le positionnement simultané réalisé dans l'infrastructure respecte les règles de la dualité. Celles-ci sont rappelées en Annexe I et nous dégagerons progressivement l'importance d'un tel critère.

En schématisant les choses nous montrerons que les méthodes françaises, et les méthodes de type INDSCAL, donnent une infrastructure inexacte mais significative, que la méthode procuste et la méthode de "superposition" offrent une infrastructure éventuellement exacte mais peu significative. L'Analyse Factorielle conJointe (AFJ) que nous proposons peut alors présentée comme une synthèse de ces méthodes donnant, sous certaines conditions, une infrastructure exacte et significative.

Le chapitre I contient un bref rappel de quelques méthodes classiques d'infrastructure. Ensuite nous exposons l'AFJ dans les chapitres II et III. Le chapitre IV présente trois applications et permet de comparer les résultats de l'AFJ avec ceux des autres méthodes. Le chapitre V est consacré à l'étude des interstructures.

CHAPITRE I :

LES METHODES D'INTRASTRUCTURE

1.1.- Rappel des principales méthodes d'intrastructure

Nous présentons ici les grandes lignes des méthodes d'analyse des données non-cubiques. Nous détaillerons ultérieurement, dans la partie de notre étude qui s'en rapproche le plus, la manière dont ces méthodes construisent leur intrastructure. Nous mentionnons également quelques méthodes d'analyse des données cubiques qui seront comparées aux précédentes, ou utilisées dans leur formulation.

a) Les méthodes pour données non-cubiques

— INDSCAL et IDIOSCAL cherchent à reconstituer une famille de tableaux de similarités W_k , $k \in K$, à partir de (semi-) métriques diagonales D_k (INDSCAL) ou non diagonales M_k (IDIOSCAL) et d'un tableau individus \times caractères X ($n \times p$). Ces méthodes minimisent donc une fonction de X et D_k (ou M_k), $k \in K$, soit : $\phi(X, D_k) = \sum_k \|W_k - XD_kX'\|_{\varphi_{DD}}^2$. Divers algorithmes ont été proposés pour résoudre ce problème. On peut voir Carroll et Chang (1970), Tucker (1972), De Leeuw et Pruzanski (1978) ou Lafaye *in* SAD (1985).

— L'Analyse Canonique Généralisée (ACG) au sens de Carroll (1968) recherche les liaisons existant entre plusieurs groupes (ou paquets) de variables. Les variables canoniques z^j sont les vecteurs D -orthogonaux de l'espace des variables $F = (\mathbf{R}^n, D)$ qui maximisent successivement la somme des carrés des coefficients de corrélation multiple entre z^j et les variables du tableau X_k , soit : $\sum_k R^2(z^j, X_k)$. On verra en Annexe II que l'ACG est aussi une méthode d'intrastructure lorsqu'on interprète ses résultats dans l'espace des individus.

— STATIS (l'Hermier des Plantes : 1976, Lavit : 1985 *in* SAD, et 1988) propose une analyse conjointe des tableaux quantitatifs en trois étapes :

- Une analyse globale donne "**l'interstructure**" qui est un positionnement multidimensionnel des m tableaux pris comme objets de l'étude. Elle définit donc des proximités entre les différentes situations k , $k \in [m]$ (voir chap. V).

- Une analyse moyenne dégage avec “le compromis” une information moyenne qui synthétise les informations contenues dans les différents tableaux.

- Une analyse fine construit une infrastructure et éventuellement des trajectoires individuelles. La technique d’obtention du positionnement simultané n’étant pas clairement définie, nous essayerons d’expliquer cette infrastructure.

— L’Analyse Factorielle Multiple (AFM - Escofier et Pagès : 1983 et 1988), peut être présentée comme une ACG pour la mesure de liaison $L^2(z, X_k) = \sum_{j \in J_k} m_{kj} \text{cov}(z, X_k^j)$, lorsqu’on a des métriques diagonales $M_k = \text{diag}(m_{kj})$. Nous verrons que la méthode dégage les liaisons entre variables dans l’espace des variables F , une infrastructure dans l’espace des individus $E = (\mathbf{R}^{\Sigma p_k}, \mathcal{M})$ et une interstructure dans l’espace des opérateurs \mathbf{R}^{n^2} (voir chap. V). Elle s’attache à montrer la cohérence des résultats obtenus à ces trois niveaux.

— La technique de superposition des axes principaux d’inertie obtenus par l’ACP de chaque tableau, n’est qu’une commodité dont nous préciserons le domaine très restreint d’utilisation.

— L’Analyse Procuste construit une infrastructure de nuages relatifs aux mêmes individus $i, i \in I$, telle que la somme des carrés des distances entre individus de même indice soit minimale. Nous exposons ci-dessous une version particulière de cette méthode.

b) Les méthodes pour données cubiques

— L’ACP 3-modes généralise au cube de données X_{IJK} , les résultats obtenus dans la décomposition en valeurs singulières d’un tableau X_{IJ} (Tucker 1966). Ce problème est appelé Tucker-3 dans Kroonenberg (1983) qui donne un exposé complet de la méthode, de ses nombreux dérivés, et des algorithmes associés. Il existe une généralisation aux n -cubes de données avec l’ACP n -modes proposée par Polit (1986).

La méthode Parafac/Candecomp dérivée du problème Tucker-2 (dans lequel on ne réduit pas le mode des situations) peut rappeler l’AFJ, ce que nous évoquerons plus loin.

— Rappelons aussi que l’ACP permet d’obtenir facilement deux infrastructures, exactes ou non, pour des données cubiques $X_{IJK} = \{X_k(n \times \bar{p}), k \in [m]\}$.

- L'ACP du Tableau Conjoint Empilé (multicentré) $X_e = \begin{pmatrix} X_1 \\ \dots \\ \vdots \\ \dots \\ X_m \end{pmatrix}$ présente un positionnement simultané des m nuages définis par les tableaux (centrés) X_k , optimal –au sens de l'inertie– pour la réunion des nuages. L'intrastructure obtenue est exacte.
- L'ACP du Tableau Moyen $\bar{X} = \sum_k \rho_k X_k$, ($\rho_k > 0$ et $\sum_k \rho_k = 1$) avec projection des tableaux X_k en supplémentaires donne une autre intrastructure, non nécessairement exacte ni optimale.

L'intérêt de cette seconde méthode est d'être adaptable au cas non-cubique des deux façons présentées ci-dessous, dans lesquelles, les poids ρ_k peuvent être choisis pour retrouver les résultats des méthodes classiques.

1.2.- Les intrastructures inexactes.

a) L'AFJ1 (ou l'AFM)

Pour tout $k \in [m]$, considérons :

$\tilde{X}_k = (0 \quad \dots \quad 0 \quad X_k \quad 0 \quad \dots \quad 0)$ tableau $(n \times p)$ avec $p =$

$\sum_k p_k$ et $\mathcal{M}_{(p \times p)} = \begin{pmatrix} M_1 & & 0 \\ & \ddots & \\ 0 & & M_m \end{pmatrix}$. Alors les triplets (X_k, M_k, D) et $(\tilde{X}_k, \mathcal{M}, D)$ sont équivalents. Ce dernier triplet donne un positionnement du nuage $N_k = \boxed{X'_k}$ dans l'espace $E^p = \bigoplus E_k$ et on le note : $\tilde{N}_k = \boxed{\tilde{X}'_k}$.

Les triplets $(\tilde{X}_k, \mathcal{M}, D)$, $k \in [m]$ constituent alors un cube de données que l'on peut traiter par la méthode du Tableau Moyen. Il s'écrit :

$$\bar{X} = \sum_k \rho_k \tilde{X}_k = (\rho_1 X_1 \quad \dots \quad \rho_m X_m) = X_\rho.$$

En général, nous prendrons des poids égaux (cf chap.V). Nous verrons que le choix des poids distingue les méthodes et influence la qualité de l'intrastructure.

Dans l'ACP du tableau moyen on diagonalise $W_0 D = X_\rho \mathcal{M} X'_\rho D = \sum_k \rho_k^2 W_k D$, avec : $W_k = X_k M_k X'_k$, pour trouver les éléments propres (C_0, Λ_0) . C_0 est un tableau $(n \times r)$ ($r = \text{rang } X_\rho$). Ses lignes sont les coordonnées des n points du nuage moyen dans la base $(u_j)_{j \in [r]}$ du sous-espace **Compromis** E_C de E^p

que ce nuage moyen engendre. La \sqrt{m} -projection (projection multipliée par \sqrt{m}) des tableaux \tilde{X}_k en supplémentaires sur ce sous-espace E_C définit l'intrastructure $\hat{X}_k = \sqrt{m}\rho_k W_k D C_0 \Lambda_0^{-1}$, $k \in [m]$.

On vérifie que : $\hat{X} = \sum_k \rho_k \hat{X}_k$ est égal à $\sqrt{m}C_0 = C$ (les propriétés barycentriques sont conservées dans une projection, un changement de base et une homothétie). $\hat{X} = C$ définit un positionnement du **nuage compromis** homothétique du nuage moyen pour l'homothétie de centre \mathbf{O}_{E_P} et de rapport \sqrt{m} . Le nuage compromis est le nuage moyen des \sqrt{m} -projetés des nuages \tilde{N}_k sur E_C et C est le **tableau compromis**. L'utilisation d'une \sqrt{m} -projection sera justifiée ultérieurement. Cette méthode est appelée AFJ1. Elle est proche de l'Analyse Factorielle Multiple qui impose des poids ρ_k proportionnels à $1/\sqrt{\lambda_1^k}$ (ou une prémultiplication des tableaux par ce facteur), avec λ_1^k première valeur propre de $W_k D$.

b) L'AFJ2 (ou Statis)

L'opérateur $W_k D$ est caractéristique du triplet (X_k, M_k, D) . Il en résulte que l'ACP du triplet (W_k, D, D) redonne le positionnement des individus obtenu dans l'ACP de (X_k, M_k, D) . Pour ce dernier triplet, $W_k D$ admet les éléments propres (C_k, Λ_k) avec $C_k' D C_k = \Lambda_k$. Pour (W_k, D, D) on diagonalise $(W_k D)^2$ dont les éléments propres (B_k, A_k) avec $B_k' D B_k = A_k$, vérifient $A_k = \Lambda_k^2$ et donc $B_k = C_k \Lambda_k^{1/2}$, ou encore $\Lambda_k = A_k^{1/2}$ et $C_k = B_k \Lambda_k^{-1/2}$.

L'AFJ2 consiste à faire l'ACP du Tableau Moyen sur le cube de données défini par les triplets (W_k, D, D) , $k \in [m]$. On trouve pour triplet moyen : $(W = \sum_k \rho_k W_k, D, D)$. La diagonalisation de $(WD)^2$ donne (B, A) , puis $\Lambda = A^{1/2}$ et $C = B \Lambda^{-1/2}$, ou plus simplement, la diagonalisation de WD donne directement (C, Λ) .

Il reste à projeter les W_k en supplémentaires (dans l'AFJ2 on considère des projections et non des \sqrt{m} -projections), pour obtenir $\hat{W}_k = W_k D C \Lambda^{-1/2}$. Peut-on en déduire un positionnement simultané \hat{X}_k , $k \in K$, des individus des différents tableaux ? Si on prend le positionnement simultané \hat{W}_k , $k \in K$ et qu'on lui applique la règle de passage de B (qui définit le positionnement \hat{W} de $\bar{W} = W$) à $C = B \Lambda^{-1/2}$ (qui définit le positionnement \hat{X} d'un certain \bar{X}), on obtient : $\hat{X}_k = W_k D C \Lambda^{-1/2} \Lambda^{-1/2} = W_k D C \Lambda^{-1}$, $k \in [m]$, et le compromis

$\hat{X} = \sum_k \rho_k \hat{X}_k = C$. Alors C désigne le tableau des composantes principales de $X_{\sqrt{\rho}} = (\sqrt{\rho_1}X_1 \quad \dots \quad \sqrt{\rho_m}X_m)$ considéré comme un tableau moyen \bar{X} .

On peut présenter ainsi l'intrastructure de la méthode Statis qui utilise des "poids" particuliers vérifiant $\sum_k \rho_k^2 = 1$.

Remarques :

- 1) Ces deux intrastructures sont inexactes sauf dans de rares cas particuliers que nous verrons plus loin.
- 2) Une isométrie transformant n'importe quel tableau X_k laisse ces intrastructures inchangées puisque $W_k = X_k M_k X_k'$ est invariant dans une isométrie.
- 3) Les appellations AFJ1 et AFJ2 peuvent surprendre puisque ces deux méthodes proposent des intrastructures très proches de celles de l'AFM ou de Statis. Nous les utilisons pour les raisons suivantes :
 - L'AFM ou Statis ne se réduisent pas à la recherche d'une intrastructure.
 - La pondération de l'AFM et celle de Statis sont fixées alors qu'elle est libre (à ce stade) dans l'AFJ1 et l'AFJ2.
 - Les intrastructures de l'AFJ1 et de l'AFJ2 sont parfaitement identiques pour des poids égaux (d'où la ressemblance des noms). Comme nous utilisons l'AFJ1 pour initialiser l'algorithme de l'AFJ et qu'elle fournit un critère permettant d'apprécier sa faisabilité, nous considérons l'AFJ1 comme la première étape de l'AFJ.

1.3.- Les intrastructures exactes.

a) La "superposition des axes principaux d'inertie".

Notons $q = \sup_k q_k$ et considérons les tableaux $C_k^c(n \times q)$ obtenus en complétant chaque tableau $C_k(n \times q_k)$ par $q - q_k$ colonnes de 0. Pour tout k , les triplets (C_k^c, I_q, D) et (X_k, M_k, D) sont équivalents. Donc les C_k^c définissent une intrastructure exacte dans l'espace (\mathbf{R}^q, I_q) obtenue par superposition des axes principaux d'inertie de même rang. Si on remplace les C_k^c par les tableaux C_k^s formés des s premières composantes, cette intrastructure limitée à s composantes conserve le maximum d'inertie. Cette propriété mise à part, il n'y a aucune raison pour superposer les axes de même rang des différents nuages. Ainsi on peut imaginer deux nuages du plan tels que le critère "Procrustes" (ou Procuste)

conduise à superposer les axes 1 et 2 et non les axes 1 comme le veut le critère de “superposition”. En fait ces deux infrastructures ne sont pas nécessairement significatives du point de vue du critère basé sur la dualité que nous construirons.

b) L'Analyse Procuste

Dans Ten Bergé et Knol (1984) le problème (orthogonal) procuste est ainsi posé : ayant $n \geq p_1 \geq p_2 \geq \dots \geq p_m$, on cherche les matrices T_k de dimensions $(p_k \times p_m)$ réalisant $\text{Min } f(T_1, T_2, \dots, T_m) = \sum_{k < \ell} \|X_k T_k - X_\ell T_\ell\|_{\varphi_{DI}}^2$, avec $T_k' T_k = I_{p_m}$, $k \in [m]$. Ce problème dit “asymétrique” diffère du problème “symétrique”, pour lequel $p_k = \bar{p}$, $k \in [m]$ (données cubiques par exemple), par le fait essentiel que les matrices T_k ne définissent pas des isométries de (\mathbf{R}^{p_k}, I) dans (\mathbf{R}^{p_m}, I) . Ainsi on trouve une infrastructure exacte pour l'analyse orthogonale procuste généralisée symétrique, mais inexacte dans le cas asymétrique. En revanche, si on choisit des matrices $T_k(p_k \times p_1)$ définissant des isométries de \mathbf{R}^{p_k} dans \mathbf{R}^{p_1} , on a $T_k T_k' = I_{p_1}$ et le cas asymétrique équivaut au cas symétrique obtenu en complétant chaque X_k par $p_1 - p_k$ colonnes de 0.

Il existe donc bien deux infrastructures exactes pour des données non-cubiques. L'une s'obtient par la méthode de “superposition” et l'autre par l'Analyse Procuste “symétrisée” présentée ci-dessus. Nous verrons qu'elles ne sont significatives que dans de rares cas particuliers.

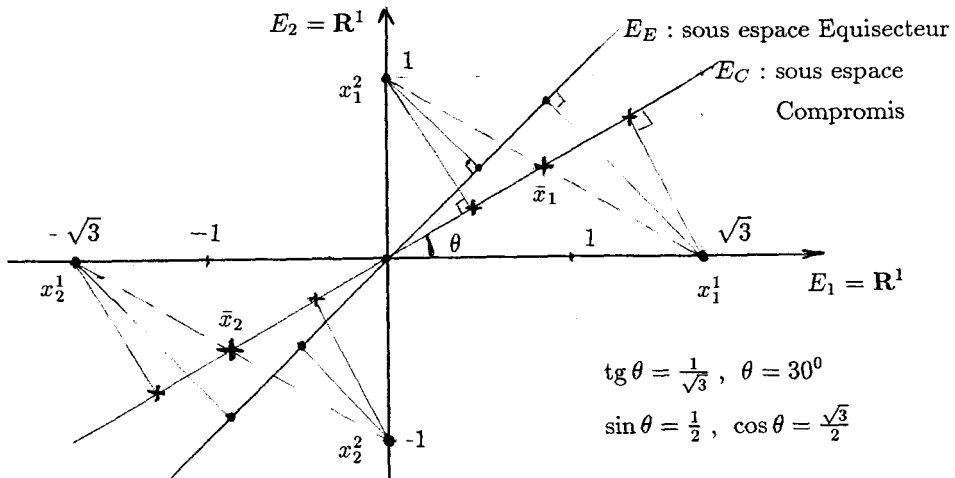
1.4.- L'AFJ et les autres méthodes.

a) Les méthodes d'infrastructure.

Dans les chapitres II et III, nous allons développer l'Analyse Factorielle conJointe – mais aussi l'ensemble des méthodes d'infrastructure – à partir de considérations géométriques qui généralisent la méthode dite “AFJ1”. Celle-ci se résume très simplement : les nuages N_k sont positionnés dans l'espace (E^p, \mathcal{M}) somme directe orthogonale des espaces (E_k, M_k) . Dans l'AFJ1 l'infrastructure s'obtient en projetant les nuages \tilde{N}_k sur le sous-espace Compromis E_C et elle est nommée Infrastructure Compromis (I.C.). De même, l'AFJ est une méthode qui obtient une infrastructure exacte en projetant sur un sous-espace particulier de E^p appelé sous-espace Equisecteur et donnant l'Infrastructure Equisecteur (I.E.). Nous préciserons cette première définition de l'AFJ, puis nous définirons dans E^p le sous-espace Déformant donnant l'Infrastructure Déformante (I.D.), et nous évoquerons les sous-espaces procuste (cf 3.3.c), discriminant (cf Annexe II),

de superposition, donnant les infrastructures correspondantes. Cette présentation établit donc un lien très simple entre toutes ces méthodes d'infrastructure.

Le petit exemple suivant peut être représenté dans le plan ; il aide à bien comprendre l'exposé du problème dans le cas général. Deux paquets de une variable ont été observés sur deux individus, et on a : $X_1 = \begin{pmatrix} \sqrt{3} \\ -\sqrt{3} \end{pmatrix}$ et $X_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. L'AFJ1 avec des poids égaux donne le tableau moyen $X_\rho = \frac{1}{2} \begin{pmatrix} \sqrt{3} & 1 \\ -\sqrt{3} & -1 \end{pmatrix}$, et les individus moyens engendrent le sous espace Compromis E_C que l'on représente dans $\mathbf{R}^2 = \mathbf{R} \oplus \mathbf{R}$ comme suit :



Les projections des nuages \tilde{N}_k sur E_C sont définies respectivement par : $\frac{\sqrt{3}}{2}X_1$ et $\frac{1}{2}X_2$, donc les \sqrt{m} -projections sont : $\hat{X}_1 = \frac{\sqrt{6}}{2}X_1$ et $\hat{X}_2 = \frac{\sqrt{2}}{2}X_2$. Cette Infrastructure Compromis est inexacte et aucune homothétie ne peut la rendre exacte.

Les projections des nuages \tilde{N}_k sur le sous espace Equisecteur E_E (première bissectrice) s'écrivent : $\frac{1}{\sqrt{2}}X_k, k \in K$, donc les \sqrt{m} -projections sont : $\hat{X}_k = X_k$. Cette Infrastructure Equisecteur est exacte.

b) Les méthodes de reconstitution des données.

Nous établirons cette deuxième définition : l'AFJ est une méthode qui recherche des triplets $(Y_k(n \times q), I, D)$ équivalents à $(X_k, M_k, D), k \in K$, des

matrices diagonales A_k ($q \times q$) et un tableau Y de dimensions ($n \times q$) tels que : $\Phi(Y_k, A_k, Y) = \sum_k \|Y_k - YA_k\|_{\varphi_{DI}}^2$ soit minimale. Cette formulation est voisine de celle des problèmes INDSCAL et éventuellement Candecomp (Carroll et Chang 1970) ou Parafac (Harshman 1970, 1984). Avec Kiers (1987, 1988), remarquons que Parafac et Candecomp cherchent une reconstitution des données sous la forme $\hat{X}_k = YA_kT'$, $k \in K$, telle que $\sum_k \|X_k - \hat{X}_k\|_{\varphi_{DM}}^2$ soit minimale, la dimension ($r \times r$) de A_k diagonale étant choisie. Pour sa part, INDSCAL recherche $\hat{W}_k = YA_k^2Y'$, $k \in K$, tels que $\sum_k \|W_k - \hat{W}_k\|_{\varphi_{DD}}^2$ soit minimale.

L'originalité de l'AFJ, qui a notre sens justifie qu'on lui donne un nom, ne tient pas tant au problème d'optimisation posé, qu'à la manière géométrique de l'aborder et de lui donner une solution :

- 1) Nous ne cherchons pas la meilleure reconstitution des données (avec \hat{X}_k ou \hat{W}_k) mais une reconstitution exacte des données (avec Y_k) permettant de plus, d'avoir une intrastucture significative.
- 2) Nous établissons des liaisons simples entre toutes les méthodes précédemment citées.
- 3) Nous contruisons un algorithme qui prouve l'existence de minima pour Φ , avec une initialisation qui donnera, en général, une convergence rapide vers un éventuel optimum global.

En proposant une double définition de l'AFJ, l'une géométrique, l'autre analytique, nous alourdissons nécessairement notre exposé. L'utilisation de la seule définition analytique permettrait de réduire l'exposé de la méthode à quelques pages (la proposition 11 du 3.2.b. et l'algorithme qui en découle en 3.2.c.). Il nous faut donc motiver l'autre définition et les considérations géométriques parfois délicates qui sont développées au chapitre II.

Leur principale raison d'être, d'ordre "historique", tient au fait que l'égalité (1) de la proposition 11 a été imaginée dans ce contexte géométrique. Cette raison historique, qui a un intérêt momentané, se double d'une raison "géographique" plus permanente : la géométrisation du problème permet de comparer simplement les différentes méthodes (françaises ou anglo-saxonnes) et de choisir une initialisation très performante de l'algorithme.

CHAPITRE II :

LES TABLEAUX A COMPOSANTES PROPORTIONNELLES

Précisons le contenu de ce chapitre II :

- En 2.1 on définit les concepts principaux de l'AFJ. Une application aux données cubiques est proposée. Pour les données non-cubiques, une infrastructure exacte et parfaitement significative n'existe que sous l'hypothèse de tableaux à composantes proportionnelles (elle donne son titre au chapitre II).

- Sous cette hypothèse, on compare en 2.2, les infrastructures de différentes méthodes. Pour obtenir des résultats précis, nous partons d'hypothèses encore plus sévères qui sont progressivement relâchées. Il s'agit d'une introduction à l'AFJ car :

1) on montre que seule son infrastructure est exacte et significative.

2) le Repère Déformant introduit dans cette démonstration, est un concept de base pour la construction de la méthode dans le cas général, faite au chapitre III.

2.1.- Le Repère Equisecteur.

a) Les systèmes équisecteurs.

Définition 1.- Soit $(e^k)_{k \in [m]}$ une base orthonormale de l'espace Euclidien $E^m = (\mathbf{R}^m, I_m)$. On appelle **vecteur équisecteur** de la base (e^k) le vecteur unitaire de E^m défini par : $e = \frac{1}{\sqrt{m}} \sum_{k \in [m]} e^k$.

De même Δe est la **direction équisectrice** des directions Δe^k .

Remarques

- Cette définition généralise à \mathbf{R}^n la notion de bissectrice.
- Dans l'espace des variables F , si $\mathbf{1}_n$ désigne le vecteurs de coordonnées égales à 1, l'axe $\Delta \mathbf{1}_n$ est la direction équisectrice des Δf^i et la D -projection d'une variable sur $\Delta \mathbf{1}_n$ définit sa moyenne.

- Les propriétés de ces objets vont être exploitées dans la généralisation suivante de la définition 1.

Définition 2.- Soit $E = (\mathbf{R}^{mq}, I_{mq})$ l'espace Euclidien somme directe orthogonale d'une famille de m espaces Euclidiens de même dimension $E_k = (\mathbf{R}^q, I_q)$, $k \in [m]$. Une base orthonormale $(e_j^k)_{j \in [q]}$ de l'espace E_k plongé dans E , sera notée $(\tilde{e}_j^k)_{j \in [q]}$. On appelle **système équisecteur** de E le système des vecteurs e_j , $j \in [q]$, équisecteurs des $(\tilde{e}_j^k)_{k \in [m]}$; soit : $\{e_j = \frac{1}{\sqrt{m}} \sum_{k \in [m]} \tilde{e}_j^k, j \in [q]\}$.

Remarques :

- Le système équisecteur $(e_j)_{j \in [q]}$ engendre un sous-espace de E noté E_e .
- Le système de vecteurs $(\tilde{e}_j^k)_{j \in [q], k \in [m]}$ constitue une base orthonormale de E .
- Il faut bien noter que chaque Δe_j est direction équisectrice d'axes $\Delta \tilde{e}_j^k$ situés dans des espaces E_k différents (ils sont supplémentaires dans E). Il n'y a donc pas d'analogie avec les propriétés de $\Delta \mathbf{1}_n$ visibles dans F .

Nous pouvons alors énoncer la proposition suivante :

Proposition 1.-

- 1) $(e_j)_{j \in [q]}$ est une base orthonormale de E_e .
- 2) Un vecteur quelconque x de E_k est caractérisé par un q -uplet de coordonnées. \tilde{x} désigne ce vecteur plongé dans E et \hat{x} sa \sqrt{m} -projection sur le sous-espace E_e . Dans la base (e_j) de E_e on a $\hat{x} \approx x$ (où \approx signifie que les deux vecteurs ont les mêmes coordonnées).
- 3) Le système équisecteur, ainsi que le sous-espace E_e dépendent de l'ordre dans lequel sont rangés les vecteurs de chacune des bases $(e_j^k)_{j \in [q]}$, $k \in [m]$.

- 1) E étant muni de la métrique $I_{mq} = I$, on a d'après la remarque précédente :

$$\forall (i, j) \in [q]^2: e_i' I e_j = \frac{1}{m} \left(\sum_k \tilde{e}_i^k \right)' I \left(\sum_\ell \tilde{e}_j^\ell \right) = \frac{1}{m} \sum_k (\tilde{e}_i^k)' I \tilde{e}_j^k = \delta_{ij}$$

- 2) $\forall k \in [m], \forall x \in E_k$, on a : $x = \sum_{j=1}^q x^j e_j^k$, et $\tilde{x} \in \tilde{E}_k$ s'écrit $\tilde{x} = \sum_j x^j \tilde{e}_j^k$.

Sa \sqrt{m} -projection sur Δe_i s'écrit : $\hat{x}^i = \sqrt{m} e_i' I \left(\sum_j x^j \tilde{e}_j^k \right)$, $i \in [q]$, soit :

$$\hat{x}^i = \sqrt{m} \frac{1}{\sqrt{m}} \left(\sum_\ell \tilde{e}_i^\ell \right)' I \sum_j x^j \tilde{e}_j^k = \sum_\ell \sum_j ((\tilde{e}_i^\ell)' I \tilde{e}_j^k) x^j$$

$$\hat{x}^i = x^i, i \in [q] \quad \text{c'est} \quad \hat{x} \approx x$$

\hat{x} est la \sqrt{m} -projection de \tilde{x} sur le sous-espace E_e .

- 3) L'ordre dans lequel sont rangés les vecteurs d'une base est arbitraire et on peut construire un système équisecteur et un sous-espace E_e pour chacune des $(q!)^{m-1}$ façons de choisir, sans ordre, q m -uplets de vecteurs du type : $(e_{j_1}^1, e_{j_2}^2, \dots, e_{j_m}^m)$, chacun donnant un vecteur e_j . Ainsi dans $\mathbf{R}^4 = \mathbf{R}^2 \oplus \mathbf{R}^2$, on a la base :

$$\tilde{e}_1^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_2^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tilde{e}_1^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tilde{e}_2^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

qui donne deux systèmes équisecteurs :

- $e_1 = \frac{1}{\sqrt{2}}(\tilde{e}_1^1 + \tilde{e}_1^2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ et $e_2 = \frac{1}{\sqrt{2}}(\tilde{e}_2^1 + \tilde{e}_2^2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$
- $f_1 = \frac{1}{\sqrt{2}}(\tilde{e}_1^1 + \tilde{e}_2^2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ et $f_2 = \frac{1}{\sqrt{2}}(\tilde{e}_2^1 + \tilde{e}_1^2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$

Ils engendrent deux sous-espaces E_e différents.

Le lecteur pourra vérifier qu'un changement des bases modifie le sous-espace E_e sauf si les matrices de changement de base sont identiques dans tous les espaces E_k . ■

b) Signification statistique pour les données cubiques.

Soit un cube de données X_{IJK} . Dans chaque espace $E_k = (\mathbf{R}_{(e_j^k)}^{\bar{p}}, I_{\bar{p}})$, X_{IJK} définit un nuage N_k , noté \tilde{N}_k lorsqu'on le plonge dans $E = \bigoplus E_k = (\mathbf{R}^{m\bar{p}}, I_{m\bar{p}})$ rapporté à la base $(\tilde{e}_j^k)_{j \in [\bar{p}], k \in [m]}$ (pour ce cas particulier, la dimension q des espaces E_k du cas général, devient \bar{p} qui est le nombre de variables de chaque tableau). Considérons un système équisecteur quelconque : $\{e_j = \frac{1}{\sqrt{m}} \sum_k \tilde{e}_j^k, j \in [\bar{p}]\}$.

Corollaire

- 1) les \sqrt{m} -projections \hat{N}_k des nuages \tilde{N}_k sur un sous-espace E_e quelconque définissent une intrastucture exacte. Celle-ci coïncide avec celle qu'on obtiendrait dans $\mathbf{R}^{\bar{p}}$ par superposition des axes $\Delta e_j^k, k \in K$, en les associant comme les \tilde{e}_j^k le sont dans le calcul de e_j

2) Il y a $(\bar{p}!)^{m-1}$ sous-espace E_e différents et autant d'intrastructures exactes.

■ 1) Chaque intrastructure est exacte car les images euclidiennes suivantes sont équivalentes : $(N_k, E_k, I_{\bar{p}}) \iff (\tilde{N}_k, E, I_{m\bar{p}}) \iff (\hat{N}_k, E_e, I_{\bar{p}})$. La première équivalence a été vue avec l'AFJ1 et la seconde résulte de la proposition précédente.

2) Le dénombrement $(\bar{p}!)^{m-1}$ déjà calculé est aussi le nombre de permutations des indices $j \in J$, à la puissance m car il y a m ensembles J ; divisé par $\bar{p}!$ car les permutations des e_j de E_e définissent le même sous-espace. ■

Proposition 2.-

- 1) Il existe un système équisecteur qui possède un sens statistique : c'est celui qui, pour tout $j \in J$, associe dans le calcul des e_j , les axes Δe_j^k , $k \in K$, liés par dualité à la même variable j (indice j attaché à une seule variable).
- 2) L'intrastructure obtenue coïncide avec le positionnement simultané des nuages N_k réalisé dans $\mathbf{R}^{\bar{p}}$ par le Tableau Conjoint Empilé X_e .

■ 1) Le corollaire précédent précise que l'intrastructure obtenue par \sqrt{m} -projection sur E_e s'obtient également en superposant les axes Δe_j^k qui sont associés dans le calcul des e_j . Une intrastructure obtenue par \sqrt{m} -projection sur un sous-espace E_e quelconque a une existence géométrique mais, en général, aucun sens statistique puisqu'elle superpose sur un seul axe, des axes relatifs à des variables différentes (sauf cas très particuliers). En effet, la dualité stipule qu'un axe quelconque de l'espace des individus est lié à une variable **déterminée** (voir Annexe I). Seul le cas mentionné ci-dessus permet une interprétation statistique puisqu'il associe dans le calcul de chaque e_j , et donc superpose, des axes liés à la même variable statistique $j \in J$.

2) Le Tableau Conjoint Empilé $X_e = \begin{pmatrix} X_1 \\ \dots \\ \vdots \\ \dots \\ X_m \end{pmatrix}$ définit le nuage $N = \bigcup_{k \in K} N_k$

qui réalise un positionnement simultané dans $\mathbf{R}^{\bar{p}}$ des nuages N_k . Celui-ci s'obtient également en superposant les axes Δe_j^k associés pour tout k , $k \in K$, à une même variable j , $j \in J$. ■

Définition 3.- Ce système équisecteur qui possède un sens statistique sera nommé **Repère Equisecteur de E** , et le sous-espace qu'il engendre s'appellera

sous-espace Equisecteur de E , notée E_E .

Remarque : Pour les données cubiques, nous venons de voir deux façons d'obtenir une représentation simultanée des différents nuages : prendre le tableau X_e ou \sqrt{m} -projeter sur E_E . Ces deux méthodes sont validées par la dualité puisqu'elles consistent à superposer des axes liés à une même variable statistique. La dualité nous permet aussi de définir une instructure de données non-cubiques.

c) Le Repère Equisecteur pour les données non-cubiques.

Définition 4.- Dans l'espace choisi pour représenter simultanément les nuages d'individus, l'instructure qui respecte les propriétés de dualité s'obtient en superposant les axes des différents nuages liés (par dualité) à une même variable ou plus généralement à une même direction de l'espace des variables. Une telle instructure est dite **significative**.

Dans cette définition, le premier cas envisagé (axes associés à une même variable j) correspond à l'instructure des données cubiques. Celle-ci sera maintenant considérée comme l'instructure de la méthode du Tableau Conjoint Empilé. Le second cas définit l'instructure de données non-cubiques dans le cas de composantes (combinaisons linéaires de variables initiales) proportionnelles. En effet des composantes proportionnelles occupent une même direction de l'espace des variables $F = (\mathbf{R}^n, D)$. Nous allons donc chercher des triplets (Y_k, I, D) équivalents aux triplets initiaux (X_k, M_k, D) vérifiant, si possible, cette propriété. Soit $q_k = \text{rang} X_k$ avec $X_k (n \times p_k)$. Si $q_k < p_k$, il est préférable de partir des triplets équivalents (C_k, I, D) avec $C_k (n \times q_k)$ définissant un positionnement des individus dans l'espace $E_k^{q_k} = (\mathbf{R}_{(u_k)}^{q_k}, I_{q_k})$, $k \in K$.

Définition 5.- La famille de triplets (X_k, M_k, D) est à **composantes proportionnelles** si pour tout k , $k \in K$, il existe une isométrie H_k de $E_k^{q_k}$ dans $E_k^q = (\mathbf{R}^q, I_q)$ avec $q = \sup_k(q_k)$, telle que $Y_k = C_k H_k'$ vérifie : $Y_k = Y A_k$ avec A_k diagonale ($q \times q$) et $Y (n \times q)$.

Nous dirons plus simplement : tableaux à composantes proportionnelles. De tels tableaux vérifient le modèle INDSCAL avec des semi-métriques car $W_k = Y A_k^2 Y'$.

Remarque : Si $p_k = \dim E_k \leq q$, on peut compléter l'espace E_k par $q - p_k$ directions telles que l'espace obtenu E_k^c soit égal à E_k^q . Sinon on pose : $E_k^c = E_k$. On a ainsi : $E_k \subset E_k^q = E_k^c$ ou $E_k^q \subset E_k = E_k^c$, $k \in K$.

Proposition 3.-

Pour une famille de tableaux à composantes proportionnelles :

- 1) Il existe un Repère Equisecteur dans $E^{p^c} = \bigoplus^{\perp} E_k^c$.
- 2) En considérant les \sqrt{m} -projections des nuages \tilde{N}_k sur le sous-espace Equisecteur E_E de E^{p^c} , l'AFJ définit une infrastructure exacte et significative appelée **Intrastructure Equisecteur**. Elle s'écrit : $\hat{Y}_k = Y_k$, $k \in K$, et pour le nuage compromis : $\bar{Y} = \sum_k \rho_k Y_k = Y \sum_k \rho_k A_k$.

■ 1) Il suffit de remarquer l'interprétation géométrique suivante, particulièrement simple si $M_k = I_{p_k}$:

Le tableau $X_k(n \times p_k)$ positionne le nuage N_k dans $E_k = (\mathbf{R}_{(e_j^k)}^{p_k}, I_{p_k})$, $C_k(n \times q_k)$ définit le même nuage par ses coordonnées dans le sous-espace $E_k^{q_k}$ de $E_k \subset E_k^c$.

$Y_k(n \times q)$ définit le même nuage par ses coordonnées dans le sous-espace E_k^q de E_k^c avec $E_k^q = (\mathbf{R}_{(v_j^k)}^q, I_q)$.

Ce sont des changements de base orthonormale dans E_k^c avec variation du nombre des composantes nécessaires à restituer N_k dans E_k^c . Alors l'espace $E^{mq} = \bigoplus^{\perp} E_k^q$ est un sous-espace de $E^{p^c} = \bigoplus^{\perp} E_k^c$ et le Repère Equisecteur de E^{mq} est aussi le Repère Equisecteur de E^{p^c} .

2) D'après les définitions et propositions précédentes, les \sqrt{m} -projections des nuages \tilde{N}_k positionnés dans E^{mq} , sur le sous-espace Equisecteur E_E définissent une infrastructure exacte et significative. $E^{mq} \subset E^{p^c}$ donc l'AFJ est bien une méthode de \sqrt{m} -projection des nuages \tilde{N}_k sur le sous-espace Equisecteur E_E de E^{p^c} . ■

Remarques :

- 1) Si un espace E_k a été complété à la dimension q , on distingue $E^{p^c} = \bigoplus^{\perp} E_k^c$ et $E^p = \bigoplus^{\perp} E_k$ pour constater que :

– le sous-espace Compromis E_C vérifie : $E_C \subset E^p \subset E^{p^c}$,

- le sous-espace Equisecteur E_E vérifie : $E_E \subset E^{m_q} \subset E^{p_c}$ mais pas : $E_E \subset E^p$.
Donc la comparaison de méthodes qui projettent les nuages \tilde{N}_k sur E_C (AFJ1) ou sur E_E (AFJ) peut se faire dans E^{p_c} et non dans E^p . (L'hypothèse $p_k \leq q$ ou rang $X_k = p_k$ entraîne $E^{p_c} = E^{m_q}$).
- l'exposé de l'AFJ seule se fait dans E^{m_q} . C'est la comparaison des méthodes qui impose d'introduire E^p et E^{p_c} .
- 2) Le sous-espace E_E de E^{p_c} dépend des tableaux X_k , en particulier E^{p_c} dépend des dimensions p_k et des rangs q_k de ces tableaux.
- 3) L'AFJ déterminera l'Intrastructure Equisecteur (I.E.) sans préciser la position de E_E sauf dans les rares cas particuliers où il coïncide avec E_C .

2.2.- Introduction à l'AFJ.

Avant d'exposer au chapitre III, la façon dont l'AFJ détermine l'I.E. pour les tableaux à composantes proportionnelles (ou même dans le cas général) nous allons vérifier que les méthodes usuelles ne donnent pas des intrastructures exactes et significatives. En partant d'hypothèses très strictes, qui s'élargiront progressivement nous montrerons les lacunes de ces intrastructures.

a) Les tableaux proportionnels.

Définition 6.- *La famille de triplets (X_k, M_k, D) constitue des tableaux de composantes proportionnels (ou tableaux proportionnels) si les tableaux des composantes principales vérifient : $C_k = a_k C$ avec $a_k \in \mathbf{R}_+^*$, $k \in K$ et C ($n \times q$).*

Sous cette hypothèse (opérateurs proportionnels) les nuages sont homothétiques donc très simples à comparer. Pourtant, nous pouvons montrer que certaines méthodes proposent des intrastructures inexactes !

Proposition 4.-

Pour les tableaux proportionnels qui vérifient $C_k = a_k C$, $k \in K$:

- 1) *l'AFJ1 à poids égaux et Statis définissent la même intrastructure inexacte.*
- 2) *l'AFJ1 avec les poids $\rho_k = \frac{1}{m a_k}$ (AFM à un facteur près), l'analyse procuste et la superposition donnent l'Intrastructure Equisecteur $\hat{C}_k = C_k$, $k \in K$.*

■ Explicitons les calculs relatifs à l'AFJ1 puis l'AFJ2

• Pour l'AFJ1 on a : $W_0 = \sum_k \rho_k^2 W_k = \sum_k \rho_k^2 a_k^2 W$ avec $W = CIC'$.

Adoptons une normalisation des a_k et de C telle que $\sum_k \rho_k^2 a_k^2 = \frac{1}{m}$.

(C, Λ) sont les éléments propres de WD , avec $C'DC = \Lambda$, et (C_0, Λ_0) ceux de W_0D , avec $C'_0DC_0 = \Lambda_0$. Comme $W_0D = \frac{1}{m}WD$, C et C_0 sont proportionnels, donc : $\frac{1}{m}WDC = C\Lambda_0 = \frac{1}{m}C\Lambda$, et : $\Lambda = m\Lambda_0$, $C = \sqrt{m}C_0$.

L'introduction du coefficient multiplicateur \sqrt{m} dans l'Intrastructure Compromis (I.C.) de l'AFJ1 est ainsi doublement justifiée :

- L'I.E. retient des \sqrt{m} -projections.
- En considérant les \sqrt{m} -projections sur le sous-espace compromis E_C , si $C_k = a_k C$ avec $\sum_k \rho_k^2 a_k^2 = \frac{1}{m}$, le tableau moyen C_0 obtenu par diagonalisation de W_0D est remplacé par le tableau compromis : $\hat{X} = C = \sqrt{m}C_0$.

Pour des tableaux proportionnels l'I.C. s'écrit donc :

$$\begin{aligned}\hat{X}_k &= \sqrt{m}\rho_k W_k D C_0 \Lambda_0^{-1} \\ &= \sqrt{m}\rho_k a_k^2 W \frac{C}{\sqrt{m}} m \Lambda^{-1} = m\rho_k a_k^2 C, \quad k \in K \\ \text{et} \quad \hat{X} &= C = \sqrt{m}C_0.\end{aligned}$$

• Pour l'AFJ2 on a $W_{II} = \sum_k \rho_k W_k$, et les "poids" ρ_k de Statist sont proportionnels aux coefficients a_k^2 : $\rho_k = \frac{a_k^2}{\beta}$, $k \in K$. Ils vérifient $\sum_k \rho_k^2 = 1$, donc $\sum_k a_k^4 = \beta^2$. Comme $W_k = a_k^2 W$, les éléments propres (C_{II}, Λ_{II}) de $W_{II}D = \sum_k \frac{a_k^2}{\beta} a_k^2 W D = \beta W D$, vérifient : $\Lambda_{II} = \beta\Lambda$ et $C_{II} = \sqrt{\beta}C$.

Statist normalise W_{II} pour que : $\|W_{II}^*\| = \frac{1}{m} \sum_k \|W_k\|$, $\|W_{II}^*\| = \frac{1}{m} \sum_k a_k^2 \|W\|$ et la normalisation $\sum_k a_k^2 = m$ conduit à : $W_{II}^* = W$ (et non $W_{II} = \beta W$). L'intrastructure s'écrit : $\hat{X} = C$ et $\hat{X}_k = W_k D C \Lambda^{-1} = a_k^2 C$, $k \in K$.

On a donc les résultats suivants :

1) Si $C_k = a_k C$, $k \in K$ avec $\sum_k a_k^2 = m$, l'intrastructure de Statist s'écrit : $\hat{X} = C$ et $\hat{X}_k = a_k C_k$, $k \in K$ (cas de l'interstructure non normée).

Dans l'AFJ1 à poids égaux, $\rho_k = \frac{1}{m}$ et $\sum_k \rho_k^2 a_k^2 = \frac{1}{m}$ entraînent $\sum_k a_k^2 = m$ d'où l'Intrastructure Compromis : $\hat{X} = C$ et $\hat{X}_k = a_k C_k$, $k \in K$.

Ces deux intrastructures identiques sont inexactes car elles grossissent les gros nuages ($a_k > 1$) et diminuent les petits. Mais elles sont significatives car elles superposent des axes (principaux) de l'espace des individus liés par dualité à des variables proportionnelles.

2) Pour les poids $\rho_k = \frac{1}{ma_k}$ qui correspondent à la normalisation : $\sum_k \frac{1}{a_k} = m$, on a l'I.C. : $\hat{X} = C$ et $\hat{X}_k = a_k C = C_k$, $k \in K$. Elle coïncide avec l'I.E. qui s'écrit $\hat{C} = C$ et $\hat{C}_k = C_k$, $k \in K$ (Pour ne pas alourdir les notations on a gardé le symbole C quoique ce tableau ne soit pas égal à celui du 1)).

Il est évident que la superposition des axes principaux donne la même infrastructure. Pour l'analyse procuste, la propriété a été démontrée sous l'hypothèse plus large qui va suivre. ■

b) Les tableaux à composantes principales proportionnelles.

Définition 6 bis.- Les tableaux à composantes principales proportionnelles sont tels que : $C_k^q = CA_k$, avec A_k diagonale ($A_k = \text{diag}(a_{kj})$, $j \in [q]$), $k \in K$. Le tableau C_k^q ($n \times q$) est le tableau C_k ($n \times q_k$) complété par $q - q_k$ colonnes de 0, dont les colonnes sont convenablement réordonnées.

En effet la composante c_k^2 de C_k peut être proportionnelle, par exemple, à c^1 . Elle passe alors en première colonne de C_k^q . Si $q_k < q$, il existe $q - q_k$ indices j , $j \in [q]$, tels que $a_{kj} = 0$ et $(C_k^q)^j = 0_{\mathbf{R}^n}$.

Les composantes principales ne sont plus globalement proportionnelles. Elles le sont individuellement. Donc les nuages ne sont plus globalement homothétiques mais ils le sont "axe par axe".

On adopte la normalisation : $\sum_k \rho_k^2 A_k^2 = \frac{1}{m} I_q$, et on suppose que $C = (c^1 \quad \vdots \quad c^2 \quad \vdots \quad \dots \quad \vdots \quad c^q)$ vérifie $\|c^1\| \geq \|c^2\| \geq \dots \geq \|c^q\|$.

Ici presque toutes les méthodes offrent une infrastructure inexacte !

Proposition 5.-

Pour des tableaux à composantes principales proportionnelles :

1) L'Infrastructure Compromis (I.C.) de l'AFJ1 : $\hat{X} = C$ et $\hat{X}_k = m\rho_k C_k^q A_k$, $k \in K$ est inexacte (elle déforme les nuages par des homothéties opérant dans les directions des axes principaux). Avec la normalisation adoptée, C est bien le tableau compromis du 1.2.a.

2) S'il y a proportionnalité de la première composante des différents nuages (si $\forall k \in K, \forall j \in [q]$, on a : $\|a_{k1} c^1\| \geq \|a_{kj} c^j\|$), alors les poids $\rho_k = \frac{1}{ma_{k1}}$ font coïncider le premier axe du nuage compromis de l'AFJ1 avec l'axe 1 du Repère Equisecteur (R.E.) de l'AFJ. Les coordonnées sur l'axe 1 de l'Infrastructure

Compromis sont exactes. C'est l'AFM à un facteur près, elle n'est donc exacte que pour la première composante de chaque nuage et sous les hypothèses considérées.

- 3) S'il y a proportionnalité des composantes principales de même rang dans les différents nuages ($\forall k \in K : \|a_{k_1} c^1\| \geq \|a_{k_2} c^2\| \geq \dots \geq \|a_{k_q} c^q\|$) alors l'intrastructure "de superposition" coïncide avec l'Intrastructure Equisecteur (I.E.) de l'AFJ. Sinon, elle est exacte mais totalement dépourvue de signification.
- 4) L'Intrastructure Procuste coïncide avec l'Intrastructure Equisecteur de l'AFJ qui s'écrit : $\hat{C}_k = C_k^q = CA_k$, $k \in K$ et $\hat{C} = \sum_k \rho_k C_k^q$.
- 5) La part d'inertie de chaque nuage conservée dans l'Intrastructure Equisecteur limitée à s axes ($s \leq q$) vaut : $p_{sk} = (\sum_{j=1}^s a_{kj}^2 \lambda_j) / (\sum_{j=1}^q a_{kj}^2 \lambda_j)$, $k \in K$. Pour l'ensemble des nuages c'est aussi la part de l'inertie conservée pour le nuage compromis : $p_s = (\sum_{j=1}^s \lambda_j) / (\sum_{j=1}^q \lambda_j)$.

■ • Certaines propositions découlent de façon évidente de la proposition 1) qui sera seule démontrée. La proposition 4) n'est pas immédiate, sa démonstration pourra être trouvée dans l'Annexe III.

• On a : $C_k^q = CA_k$, $k \in K$ avec $\sum_k \rho_k^2 a_{kj}^2 = \frac{1}{m}$, $j \in [q]$. Soient (C, Λ) les éléments propres de $WD = CC'D$, (C_k^q, Λ_k^q) ceux de $W_k D$, $k \in K$.
 $C_k^q = CA_k \Rightarrow \lambda_j^k = \|c_k^j\|^2 = a_{kj}^2 \lambda_j$, $j \in [q]$ (ou $\Lambda_k^q = \Lambda A_k^2$), $k \in K$.
 Comme C ($n \times q$) est de rang q : $a_{kj} = 0 \iff \lambda_j^k = 0$, et
 $W_k DC_k^q = C_k^q \Lambda_k^q \iff W_k DC A_k = CA_k \Lambda_k^q \iff W_k DC = C \Lambda_k^q$.
 Alors $W_0 DC = \sum_k \rho_k^2 W_k DC = C (\sum_k \rho_k^2 \Lambda_k^q) = C \sum_k \rho_k^2 A_k^2 \Lambda = C \Lambda \frac{1}{m} = C \Lambda_0$.
 Les éléments propres (C_0, Λ_0) de $W_0 D$ vérifient donc $\Lambda_0 = \Lambda/m$ et $C_0 = C/\sqrt{m}$.

• L'I.C. de tableaux à composantes **principales** proportionnelles s'écrit alors :

$$\hat{X} = \sqrt{m} C_0 = C$$

et

$$\hat{X}_k = \sqrt{m} \rho_k W_k DC_0 \Lambda_0^{-1} = \sqrt{m} \rho_k W_k D \frac{C}{\sqrt{m}} m \Lambda^{-1}$$

soit

$$\hat{X}_k = m \rho_k C A_k^2 = m \rho_k C_k^q A_k, \quad k \in K$$

• L'I.C. avec des poids égaux s'écrit : $\hat{X} = C$ et $\hat{X}_k = C_k^q A_k$, $k \in K$ et les déformations subies par les nuages sont évidentes. ■

Concluons en soulignant que sous l'hypothèse de composantes principales proportionnelles seules l'AFJ et l'analyse procuste proposent une infrastructure exacte et significative. Nous verrons que pour des composantes (quelconques) proportionnelles l'infrastructure procuste n'est plus parfaitement significative (voir proposition 12). Examinons donc le cas des composantes proportionnelles.

c) Le Repère Déformant.

Définition 7.- Soient les tableaux X_k à composantes proportionnelles avec les poids ρ_k et $(v_j^k)_{j \in [q]}$ la base de E_k^q telle que $Y_k = Y A_k$, $k \in K$. On appelle **Repère Déformant (R.D.)** de $E^{p \cdot c}$ le système de vecteurs défini par : $\{v_j = \sqrt{m} \sum_{k \in K} \rho_k a_{kj} v_j^k, j \in [q]\}$, et l'hypothèse de normalisation $\sum_k \rho_k^2 A_k^2 = \frac{1}{m} I_q$. Le R.D. engendre le sous-espace Déformant E_D et les \sqrt{m} -projections des nuages \tilde{N}_k sur celui-ci définissent l'**Infrastructure Déformante**.

Le Repère Déformant est un Repère Equisecteur pondéré. Il nous donne une expression simplifiée de l'Infrastructure Compromis qui éclaire la façon dont elle est inexacte quand les tableaux sont à composantes (quelconques) proportionnelles. Dans le cas général, l'I.D. constitue un trait d'union entre l'I.C. qui est connue, et l'I.E. que l'on recherche.

Proposition 6.-

Pour une famille de tableaux X_k à composantes proportionnelles :

- 1) $(v_j)_{j \in [q]}$ est base orthonormale de E_D et un vecteur quelconque $\tilde{y} = (y^1, \dots, y^q)$ de $E_{k(v_j^k)}^q$ a pour projection sur $E_{D(v_j)}$ le vecteur $\hat{y} = \sqrt{m}(\rho_k a_{k1} y^1, \dots, \rho_k a_{kq} y^q)$.
 - 2) L'Infrastructure Déformante s'écrit : $\hat{Y}_k = m \rho_k Y A_k^2$, $k \in K$, et le compromis $\hat{Y} = \sum_k \rho_k \hat{Y}_k$ vaut Y .
 - 3) Les sous-espaces compromis E_C et Déformant E_D sont confondus. Donc les Infrastructures Compromis (I.C.) et Déformante (I.D.) obtenues par \sqrt{m} -projection sur E_C et E_D sont identiques. Leurs expressions diffèrent car elles ne sont pas rapportées à la même base (sauf si on a les composantes principales proportionnelles).
- 1) $\forall (i, j) \in [q]^2 : v_j' I v_i = \sqrt{m}(\sum_k \rho_k a_{kj} v_j^k)' I \sqrt{m}(\sum_\ell \rho_\ell a_{\ell i} v_i^\ell)$
avec $(\tilde{v}_j^k)' I \tilde{v}_i^\ell = \begin{cases} 1 & \text{si } k = \ell \text{ et } j = i \\ 0 & \text{sinon} \end{cases}$

$$\text{d'où : } v'_j I v_i = \begin{cases} 0 & \text{si } j \neq i \\ m \sum_k \rho_k^2 a_{kj}^2 = 1 & \text{si } j = i \end{cases}$$

$$\bullet \quad \tilde{y}' I v_i = \left(\sum_{j \in [q]} y^j \tilde{v}_j^k \right)' I \sqrt{m} \sum_{\ell \in K} \rho_\ell a_{\ell i} \tilde{v}_i^\ell$$

$$\text{d'où} \quad \hat{y}^i = \tilde{y}' I v_i = \sqrt{m} \rho_k a_{ki} y^i$$

$$\text{et} \quad \hat{y} = \sqrt{m} (\rho_k a_{k1} y^1, \dots, \rho_k a_{kq} y^q)$$

Pour des poids égaux, la \sqrt{m} -projection de \tilde{y} sur E_D s'écrit :

$$\hat{y} = (a_{k1} y^1, a_{k2} y^2, \dots, a_{kq} y^q).$$

2) L'I.D. s'obtient par \sqrt{m} -projection des nuages \tilde{N}_k sur E_D et donc les formules d'intrastructure découlent immédiatement de ce qui précède :

$\hat{Y}_k = m \rho_k Y_k A_k = m \rho_k Y A_k^2$, $k \in K$, et $\hat{Y} = \sum_k \rho_k \hat{Y}_k = Y$, du fait de la normalisation : $\sum_k \rho_k^2 a_{kj}^2 = \frac{1}{m}$, $j \in [q]$.

3) Montrons d'abord $E_C \subset E_D$, ou encore que tous les individus du nuage compromis (qui constituent un système de générateurs de E_C) sont dans E_D . On raisonne dans E^{mq} rapporté à la base $(\tilde{v}_j^k)_{j \in [q], k \in K}$. Dans E_k^q on a $Y_k = Y A_k$, $k \in K$ d'où les coordonnées de l'individu \tilde{y}_i^k dans la base $(\tilde{v}_j^k)_{j \in [q], k \in K}$:

$$\tilde{y}_i^k = \sum_{j \in [q]} y_i^{jk} \tilde{v}_j^k + 0, \text{ avec } y_i^{jk} = a_{kj} y_i^j.$$

Le i -ème individu compromis \tilde{y}_i , $i \in I$, est le \sqrt{m} -barycentre des \tilde{y}_i^k , $k \in K$, et :

$$\tilde{y}_i = \sqrt{m} \sum_k \rho_k \tilde{y}_i^k = \sqrt{m} \sum_k \rho_k \sum_j a_{kj} y_i^j \tilde{v}_j^k = \sum_j y_i^j \sqrt{m} \sum_k \rho_k a_{kj} \tilde{v}_j^k = \sum_j y_i^j v_j.$$

Tous les individus compromis sont dans E_D . Comme $\text{Dim} E_C = \text{Dim} E_D = q$, on a bien $E_C = E_D$. L'I.C. et l'I.D. sont donc confondues mais l'expression de l'I.C. : $\hat{X}_k = m \rho_k W_k D C \Lambda^{-1}$ est relative à la base $(u_j)_{j \in [q]}$ portant les axes principaux du compromis (elle sera nommée Repère Compromis dans la définition 9), et l'I.D. à la base $(v_j)_{j \in [q]}$ (Repère Déformant) de $E_D = E_C$. Soit $Q = (u_j)_{(v_j)}$ matrice des coordonnées des u_j dans la base v_j . Alors $Y = C Q'$ et $\hat{X}_k = \hat{Y}_k Q$, $k \in K$. ■

Définition 8.- *L'Intrastructure Déformante exprimée dans la base liée aux axes principaux du nuage moyen sera nommée Intrastructure Déformante Principale (I.D.P.) et cette base le Repère Déformant Principal (R.D.P.) ; sous l'hypothèse considérée l'Intrastructure Compromis est identique à l'I.D.P. de même que le Repère Compromis (R.C.) coïncide avec le R.D.P. (u_j)*

Conclusion.- L'introduction du Repère Déformant nous a permis de montrer plusieurs propriétés de l'Intrastructure Compromis de tableaux à composantes proportionnelles :

- Elle est déformante, avec, si $\rho_k = \frac{1}{m}$: $\hat{Y}_k = Y_k A_k, k \in K$ et $\hat{Y} = Y$. Les déformations sont le fait d'homothéties de rapport a_{kj} et de centre 0 dans chaque direction Δv_j .
- Elle est significative puisqu'on superpose les axes Δv_j^k des différents nuages, liés par dualité à la même direction Δy^j de l'espace des variables, mais elle est inexacte.

L'Intrastructure Equisecteur de l'AFJ : $\hat{Y}_k = Y_k, k \in K$ et $\bar{Y} = \sum_k \rho_k \hat{Y}_k$ est exacte et significative car elle superpose les mêmes axes Δv_j^k . La différence est claire et bien maîtrisée : l'I.C. s'obtient en \sqrt{m} -projetant les nuages sur le sous-espace E_D engendré par le R.D., alors que l'I.E. est une \sqrt{m} -projection sur E_E engendré par le R.E.

L'appellation "Repère Déformant" ne doit pas être mal interprétée. Tous les repères (considérés comme bases de sous-espaces particuliers) sont déformants, y compris le R.E. qui multiplie les nuages par $1/\sqrt{m}$. Ce nom précise que les déformations subies par les différents nuages sont encore très simples. Pour des tableaux quelconques, les déformations peuvent être, selon les méthodes (et donc selon les sous-espaces de projection), beaucoup plus complexes.

Dans cette partie géométrique, nous avons montré la possibilité de définir une intrastructure exacte et significative de tableaux à composantes proportionnelles. Par ailleurs, nous avons établi que les méthodes classiques donnent une intrastructure inexacte ou peu significative sauf dans des conditions très restrictives. Cette démarche permet maintenant de proposer une caractérisation des tableaux à composantes proportionnelles dont la mise en œuvre algorithmique donne une solution au problème de l'AFJ, y compris dans le cas général.

CHAPITRE III :

L'AFJ D'UNE FAMILLE DE TABLEAUX QUELCONQUES

3.1.- Caractérisation des tableaux à composantes proportionnelles.

Rappelons que l'hypothèse des tableaux à composantes proportionnelles (définition 5) permet la vérification du modèle INDSCAL avec des semi-métriques. Ce modèle recherche pour une famille de tableaux $W_k (n \times n)$, un tableau $Y (n \times r)$ et des (semi)-métriques diagonales D_k tels que $W_k = Y D_k Y'$, $k \in K$. En général, le modèle INDSCAL ne se vérifie pas exactement. Les méthodes existantes recherchent alors $Y (n \times r)$ et D_k , $k \in K$, tels que $\sum_k \|W_k - Y D_k Y'\|_{\varphi_{DD}}^2$ soit minimale. Ici, nous minimiserons la fonction $\sum_k \|Y_k - Y A_k\|_{\varphi_{DI}}^2$ beaucoup plus significative pour la recherche d'une infrastructure.

En fait, nous allons travailler sur une autre fonction, qui lui est égale à une constante près, mais beaucoup plus simple à minimiser. Cette autre fonction a été trouvée en recherchant une nouvelle caractérisation des tableaux à composantes proportionnelles qui complète les caractérisations des tableaux vérifiant IDIOSCAL et INDSCAL énoncées par Glaçon (1981) ou de Leeuw et Heiser (1982). Cette caractérisation traduit certaines propriétés géométriques des Repères Compromis et Déformant.

a) Propriétés du Repère-Compromis.

Considérons les espaces : $E^p = \bigoplus E_k$, $E^{mq} = \bigoplus E_k^q$ et $E^{pc} = \bigoplus E_k^c$ définis dans la proposition 3, et reprenons dans E^p , l'AFJ1 qui définit l'I.C., sous une autre forme que celle exposée en 1.2. a).

- Dans celle-ci, on considérerait les triplets (X_k, M_k, D) et $E_{(\tilde{e}_j^k)}^p$ c'est-à-dire E^p rapporté à la base $(\tilde{e}_j^k)_{k \in K, j \in [p_k]}$.
- On prend maintenant les triplets équivalents (C_k, I_{q_k}, D) et $E_{(\tilde{u}_j^k)}^{\Sigma q_k} \subset E^p$ avec les conventions suivantes :

- $(u_j^k)_{j \in [q_k]}$ est dite la base principale du sous-espace $E_k^{q_k}$ de E_k car elle est constituée des vecteurs M_k -orthonormés qui portent les axes principaux du nuage N_k .

- elle peut être complétée pour donner $(u_j^k)_{j \in [p_k]}$ base de E_k ou $(\tilde{u}_j^k)_{j \in [q]}$ base de E_k^q . Le tableau C_k ($n \times q_k$) devient alors $C_k^0 = (C_k \ ; \ 0)$ ($n \times p_k$) ou $C_k^c = (C_k \ ; \ 0)$ ($n \times q$). Ainsi les propriétés obtenues dans $E^{\Sigma q_k}$ s'étendent à E^p et E^{mq} .

Définition 9.- On appelle **Repère Compromis (R.C.)** la base $(u_j)_{j \in [r]}$ du sous-espace compromis E_C de $E^{\Sigma q_k}$ formée des vecteurs orthonormés qui portent les axes principaux d'inertie du nuage compromis N défini par le tableau $\sqrt{m}C_\rho = \sqrt{m} \sum_k \rho_k \tilde{C}_k$, de rang r .

Cette définition crée une similitude des notations pour les sous-espaces E_E , E_D , E_C et leurs bases appelées R.E., R.D., R.C..

Rappelons que $C_0(n, r)$ désigne le tableau des composantes principales de C_ρ ou de X_ρ , et que $C = \sqrt{m}C_0$ est le tableau compromis.

Proposition 7.-

Supposons que $\rho_k = \frac{1}{m}$, $k \in K$.

- 1) Les \sqrt{m} -projections des bases principales sur le Repère Compromis sont les lignes des tableaux $P_k = C_k' D C \Lambda^{-1}$, $k \in K$.
 - 2) $C_k = C P_k'$, $k \in K$: chaque nuage N_k est le \sqrt{m} -projeté du nuage compromis N sur le sous-espace E_k .
- 1) Les triplets équivalents (X_k, M_k, D) et (C_k, I, D) définissent le même tableau W_k , $k \in K$, donc le même "compromis" $W_0 = \sum_k \rho_k^2 W_k$ et le même tableau $C(n \times r)$ des composantes du nuage compromis avec $C = \sqrt{m} \cdot C_0$.

(C_0, Λ_0) sont les éléments propres de $W_0 D$ et on a, en ACP, les coordonnées des vecteurs $(u_j)_{j \in [r]}$ avec : $U = C_\rho' D C_0 \Lambda_0^{-1}$

$$U = \frac{1}{m} \begin{pmatrix} C_1' \\ \dots \\ C_2' \\ \dots \\ \vdots \\ \dots \\ C_m' \end{pmatrix} D \frac{C}{\sqrt{m}} m \Lambda^{-1} = \frac{1}{\sqrt{m}} \begin{pmatrix} P_1 \\ \dots \\ P_2 \\ \dots \\ \vdots \\ \dots \\ P_m \end{pmatrix} \text{ avec } P_k = C_k' D C \Lambda^{-1}, k \in K.$$

Le tableau $U(\sum_k q_k \times r)$ contient en élément $(\sum_{l=1}^{k-1} q_l + i, j)$, la coordonnée du vecteur u_j sur l'axe $\Delta \tilde{u}_i^k$ qui vaut : $u_j' I \tilde{u}_i^k = (\tilde{u}_i^k)' I u_j$. C'est aussi la coordonnée du vecteur \tilde{u}_i^k sur l'axe Δu_j . On a : $P_k = \sqrt{m} \tilde{U}_k' I U$, $k \in K$. Ses lignes sont les \sqrt{m} -projections des \tilde{u}_j^k sur la base (u_j) .

Ce résultat établi dans $E^{\Sigma q_k}$ s'étend à E^p et E^{mq} en prenant C_k^0 ou C_k^c .

On l'obtient très simplement dans $E^p = (\mathbf{R}^p, \mathcal{M})$ en rappelant que $U = X_p' D C_0 \Lambda_0^{-1}$ et $\tilde{U}_k' = (0 : \dots : \Lambda_k^{-1} C_k' D X_k : \dots : 0)$, et en remarquant que :

$$\sqrt{m} \tilde{U}_k' M U = \sqrt{m} \Lambda_k^{-1} C_k' D X_k M_k \frac{1}{m} X_k' D \frac{C}{\sqrt{m}} m \Lambda^{-1} = C_k' D C \Lambda^{-1} = P_k, \quad k \in K.$$

2) Dans notre première démonstration du 1) on peut écrire la formule de reconstitution des données suivante : $C_p = (C_0 \Lambda_0^{-1/2}) \Lambda_0^{1/2} U' = \frac{1}{\sqrt{m}} C U'$, soit $\frac{1}{m} (C_1 : C_2 : \dots : C_m) = \frac{1}{m} C (P_1' : \dots : P_m')$.
Donc $C_k = C P_k'$, avec $P_k = C_k' D C \Lambda^{-1}$, $k \in K$.

Soit \hat{x} la \sqrt{m} -projection d'un x quelconque de E_C sur $E_k^{q_k}$; on a $\hat{x} = P_k x$.
En effet, $\forall x \in E_C : x = \sum_{j=1}^r x^j u_j$ et sa \sqrt{m} -projection sur $\Delta \tilde{u}_i^k$ s'écrit :
 $\hat{x}^i = \sqrt{m} (\tilde{u}_i^k)' I x = \sqrt{m} (\tilde{u}_i^k)' I \sum_j x^j u_j = (P_k)_i x$, et $\hat{x} = P_k x$.
 P_k est un \sqrt{m} -projecteur de E_C sur $E_k^{q_k}$. ■

b) Propriétés du Repère Déformant.

Le Repère Déformant nous permet de trouver ici une caractérisation des tableaux à composantes proportionnelles qui est à la base de notre méthode.

Supposons que les tableaux X_k soient à composantes proportionnelles et considérons le Repère Déformant de $E^{mq} = \bigoplus_k E_k^q$. Dans chaque espace E_k^q on dispose des bases orthonormales (\tilde{u}_j^k) et (v_j^k) et on note $Q_k = (\tilde{u}_j^k)_{(v_j^k)}$ le tableau des coordonnées des (\tilde{u}_j^k) dans la base des (v_j^k) , d'où $Y_k' = Q_k (C_k^c)'$, $k \in K$. De même dans le sous-espace Déformant ou Compromis ($E_C = E_D$) on a : $Q = (u_j)_{(v_j)}$, et $Y' = Q C' \iff Y = C Q' \iff C = Y Q$.

Proposition 8.-

Soient, pour $k \in K$, les triplets (X_k, M_k, D) équivalents à (C_k, I_{q_k}, D) , $C(n \times r)$ et les tableaux $P_k^c = (C_k^c)' D C \Lambda^{-1}$ de dimensions $(q \times r)$.

1) *IDIOSCAL* admet une solution exacte à la Condition Nécessaire et Suffisante que : $r = q = q_k$, (ou $r = q = \sup_k q_k$ si on accepte des semi-métriques).

2) *INDSCAL* admet une solution exacte à la Condition Nécessaire et Suffisante que : $r = q = q_k$ (ou $r = q = \sup_k q_k$ pour des semi-métriques) et $P_k^c = Q_k' A_k Q$ avec Q et Q_k orthogonales ($q \times q$), A_k diagonale (s.)d.p., $k \in K$. Le cas ($r = q = \sup_k q_k$) et $P_k^c = Q_k' A_k Q$ caractérise les tableaux à composantes proportionnelles.

■ 1) On utilise la proposition 7-3) :

$C_k = CP_k' \Rightarrow W_k = CP_k' P_k C' = CO_k C'$ avec $O_k(r \times r)$ s.d.p., $k \in K$. Donc IDIOSCAL admet toujours une solution exacte avec des semi-métriques. On peut exiger, en notant $q = q_0 = \sup_k q_k$, qu'au moins une matrice O_{k_0} soit une métrique. Rang $C = r \geq q_k = \text{rang } W_k = \text{rang } O_k$ entraîne la CNS indiquée.

2) CN : Pour des tableaux à composantes proportionnelles, $r = q = \sup_k q_k$, et l'I.C. coïncide avec l'I.D.P.. On a donc : $\hat{X}_k = W_k D C \Lambda^{-1} = Y_k A_k Q = C_k^c Q_k' A_k Q$, $k \in K$. Remplaçons C_k^c par I_q pour \sqrt{m} -projeter les vecteurs (\hat{u}_j^k) et non les nuages \tilde{N}_k sur le sous-espace compromis, on obtient : $P_k^c = Q_k' A_k Q$, $k \in K$. La condition A_k s.d.p. revient à choisir convenablement l'orientation des axes Δv_j^k de chaque nuage.

CS : $W_k = CP_k' P_k C'$ s'écrit sous les conditions retenues :

$$W_k = C Q' A_k Q_k Q_k' A_k Q C' = Y A_k^2 Y' \text{ et,}$$

(X_k, M_k, D) est équivalent à $(Y A_k, I, D)$, $k \in K$. ■

Cette caractérisation des tableaux à composantes proportionnelles (proposition 8-2) équivaut à la caractérisation suivante établie par Glaçon (1981) :

INDSCAL admet une solution exacte si $r = q = \sup_k q_k$, et si les $O_k(q \times q)$ solutions de $W_k = CO_k C'$ admettent un même système orthonormal de vecteurs propres.

En effet : $\forall k \in K$, $O_k = P_k' P_k = Q' A_k Q_k Q_k' A_k Q = Q' A_k^2 Q$.

Dans un contexte plus large, de Leeuw et Heiser (1982) proposent une autre caractérisation équivalente (avec comme deuxième condition : $W_l W W_k = W_k W W_l$, pour tout k, l , de K), et la proposition suivante :

Proposition 9.- (unicité)

Pour des tableaux à composantes proportionnelles, une normalisation et une orientation des axes données, l'écriture $Y_k = Y A_k$, $k \in K$ est unique à la Condition

Nécessaire et Suffisante que pour tout i et j de $[q]$, il existe $k \in K$ tel que $i \neq j$ entraîne $a_{ki} \neq a_{kj}$.

■ Pour tout $i \neq j$ de $[q]$, il existe $k \in K$ tel que les valeurs propres a_{ki}^2 et a_{kj}^2 du tableau $O_k = Q' A_k^2 Q$ soient distinctes. Donc le système Q' des vecteurs propres orthonormés est unique (au signe près). L'écriture : $W_k = C O_k C' = C Q' A_k^2 Q C' = Y A_k^2 Y'$, $k \in K$ se vérifie pour des tableaux $Y_k = Y A_k$, $k \in K$ uniques (car A_k s.d.p.). Et réciproquement. ■

Notre caractérisation est cependant plus utilisable car elle permet de vérifier l'hypothèse ou de voir jusqu'à quel point on s'en écarte. Il suffit pour cela de calculer la distance $d = \sum_k \|P_k^c - Q'_k A_k Q\|^2$. Le choix de la norme induite par φ_{IA} donnera un résultat particulièrement intéressant.

La condition nécessaire $r = q = \sup_k q_k$ mérite un commentaire : elle signifie que les variables des tableaux X_k engendrent dans \mathbf{R}^n des sous-espaces inclus dans l'un deux. Pour l'Analyse Canonique Généralisée ce cas est trivial et dans l'Intrastructure Discriminante associée (cf. Annexe II) tous les nuages sont confondus. Si en revanche on s'écarte trop de cette situation, c'est la notion d'Intrastructure Equisecteur (et d'AFJ) qui perd tout sens et l'ACG permet de voir pourquoi. Ces deux méthodes sont donc complémentaires au sens où elles concernent des familles de tableaux aux caractéristiques différentes.

3.2.- L'Analyse Factorielle conJointe.

Nous abordons maintenant le cas général avec une famille de triplets n'admettant pas des composantes exactement proportionnelles. Les résultats précédents conduisent rapidement à l'algorithme recherché.

a) Les composantes conjointes

Avec des triplets (Y_k, I, D) équivalents à (X_k, M_k, D) , on peut écrire $Y_k = Y A_k + \varepsilon_k$ avec A_k diagonale s.d.p., $\varepsilon_k(n \times q)$ et $Y(n \times q)$.

Les tableaux ε_k peuvent s'interpréter comme les erreurs dans l'approximation de Y_k par $Y A_k$, $k \in K$; et s'ils sont nuls, les tableaux sont à composantes proportionnelles ce qui se vérifie très rarement.

Définition 10.- *On appelle composantes conjointes d'une famille de triplets (X_k, M_k, D) équivalents à (Y_k, I, D) , les q ensembles de m colonnes de*

même indice j dans les tableaux Y_k , tels que :

$$\Phi(Y_k, A_k, Y) = \sum_k \|\varepsilon_k\|_{\varphi_{DI}}^2 = \sum_k \|Y_k - Y A_k\|_{\varphi_{DI}}^2$$

soit minimale. Les composantes conjointes sont donc les ensembles $\{y_k^j, k \in K\}_{j \in [q]}$ qu'on désignera simplement par composantes (des tableaux) $Y_k, k \in K$. L'AFJ est une méthode de détermination de ces composantes conjointes qui sont les plus proches de la proportionalité au sens des moindres carrés.

A ces composantes sont associés un Repère Equisecteur et un Repère Déformant et les infrastructures correspondantes. L'I.E. calculée par l'AFJ est d'autant plus significative que Φ est voisin de 0, ce que nous apprécierons à l'aide du rapport $e = (\sum_k \|\varepsilon_k\|^2) / m \|Y\|^2$. L'unicité des composantes conjointes se vérifie sauf dans le cas où la fonction Φ présente plusieurs optimums locaux de même valeur auxquels correspondent plusieurs infrastructures dont aucune ne peut être considérée comme la plus significative. La multiplicité des solutions paraît d'autant plus improbable que e est petit (cf. proposition 9). Dans le cas contraire, l'I.E. étant moins significative, elle devra s'interpréter avec d'autant plus de prudence qu'elle est susceptible d'être multiple, avec des optimums locaux égaux qui seraient obtenus avec des initialisations différentes de l'algorithme que nous allons présenter.

Proposition 10.-

1) Pour des couples $(Y_k, A_k), k \in K$ donnés, avec $\sum_k A_k^2 = m I_q, \Phi(Y_k, A_k, Y)$ est minimale si : $Y = \frac{1}{m} \sum_k Y_k A_k$.

2) Alors, si $D = \frac{1}{n} I_n$, les tableaux ε_k vérifient : $\sum_k \varepsilon_k A_k = 0$.

■ 1) $\Phi = \sum_k \|Y_k - Y A_k\|_{\varphi_{DI}}^2 = \sum_k \text{tr}(Y_k' D Y_k + Y' D Y A_k^2 - 2 Y_k' D Y A_k)$. Prenons $D = \frac{1}{n} I_n$ pour simplifier les écritures. La minimisation de Φ équivaut à la minimisation de

$$\Phi_1(Y) = \sum_{i=1}^n \sum_{j=1}^q (y_i^j)^2 \sum_k a_{kj}^2 - 2 \sum_k \sum_i \sum_j y_i^j y_{ki}^j a_{kj}.$$

Ce minimum classique s'obtient lorsque : $\frac{\partial \Phi_1}{\partial y_i^j} = 0, i \in I, j \in [q]$. Avec les hypothèses de normalisation et de poids égaux on a $\sum_k a_{kj}^2 = m, j \in [q]$, et les solutions vérifient $2m y_i^j = 2 \sum_k y_{ki}^j a_{kj}, i \in I, j \in [q]$; soit $Y = \frac{1}{m} \sum_k Y_k A_k$.

2) L'Intrastructure Déformante s'écrit : $\hat{Y}_k = Y_k A_k = Y A_k^2 + \varepsilon_k A_k, k \in K$;

et le tableau moyen : $\bar{Y} = \frac{1}{m} \sum_k \hat{Y}_k = Y \frac{1}{m} \sum_k A_k^2 + \frac{1}{m} \sum_k \varepsilon_k A_k$.

On a aussi : $\bar{Y} = \frac{1}{m} \sum_k Y_k A_k = Y$ et $\frac{1}{m} \sum_k A_k^2 = I$ d'où $\sum_k \varepsilon_k A_k = 0$.

A l'optimum, on a donc : $Y = \frac{1}{m} \sum_k Y_k A_k$ et $\sum_k \varepsilon_k A_k = 0$. ■

b) Nombre minimal admissible de composantes.

L'optimum obtenu pour Φ définit une Intrastructure (I.E.) significative sous réserve que le critère e précédent soit suffisamment petit, par exemple inférieur à $e_0 = 10$ ou 20%. On dit alors que les tableaux sont à **composantes quasi-proportionnelles**.

Il serait souhaitable d'avoir un critère permettant d'apprécier, au début des calculs, la valeur de e . On a des composantes proportionnelles à la condition nécessaire que r (le rang de C), soit égal à $q = \sup_k(q_k)$, c'est-à-dire, pour les valeurs propres de WD : $\lambda_{q+1} = \lambda_{q+2} = \dots = \lambda_n = 0$.

On admettra donc la condition nécessaire : $p_q = (\sum_{i=1}^q \lambda_i) / (\sum_{i=1}^n \lambda_i) \geq p_0$ avec $p_0 = 80$ à 90% pour qu'il existe des composantes quasi-proportionnelles.

Règle n° 1 : Soit $q = \sup_k q_k$ (éventuellement $q = \sup_k p_k$), si la part d'inertie p_q portée par les q premiers axes d'inertie du nuage compromis dépasse le seuil p_0 on peut chercher une Intrastructure Equisecteur. Sinon, cette recherche est douteuse.

Nous avons traité des cas pour lequel $p_q = 99\%$ et on peut alors se limiter à un nombre s de composantes ($s < q$) tel que $p_s = (\sum_{i=1}^s \lambda_i) / (\sum_{i=1}^n \lambda_i) \geq p_0^s$ avec $p_0^s = 95\%$ à 99%, à condition que dans chaque nuage on vérifie également $p_s^k \geq p_0^s$.

Règle n° 2 : Si $p_s \geq p_0^s$ et $p_s^k \geq p_0^s, k \in K$, on peut remplacer les tableaux $C(n \times q)$ et $C_k(n \times q)$ par les tableaux C_s et C_k^s obtenus en ne gardant que les s premières colonnes de ces tableaux. Les tableaux $Y_k(n \times s)$ définissant l'Intrastructure Equisecteur contiennent la même inertie que C_k^s et donc perdent moins de 5% de l'inertie des nuages N_k .

Les seuils p_0 et p_0^s sont donnés ici à titre indicatif. Alors que p_0 repère l'adéquation des données à un modèle hypothétique ($Y_k = Y A_k$), le seuil p_0^s exprime un abandon de précision dans le positionnement des nuages et on doit être beaucoup plus strict sur ce point.

Proposition 11.-

Soient une famille de triplets (X_k, M_k, D) , les tableaux $C_k (n \times s)$ des s premières composantes principales de chaque nuage ($s \leq q$), un tableau $C (n \times s)$ vérifiant $C'DC = \Lambda$ (Λ diagonale définie positive), et les tableaux $P_k = C'_k DC \Lambda^{-1}$, $k \in K$. Pour des matrices orthogonales Q et Q_k telles que : $Y = CQ'$ et $Y_k = C_k Q'_k$, $k \in K$, et des A_k diagonales s.d.p., on a :

$$1) \sum_k \|P_k - Q'_k A_k Q\|_{\varphi_{I\Lambda}}^2 + \delta_s = \sum_k \|Y_k - Y A_k\|_{\varphi_{DI}}^2. \quad (1)$$

2) C étant le tableau des composantes principales du nuage compromis, on a $\delta_s = 0$ si $s = q = r = \text{rang} X_\rho$.

■ 1) Rappelons que $\langle A, B \rangle_{\varphi_{M_1 M_2}} = \text{tr}(A' M_1 B M_2) = \text{tr}(B M_2 A' M_1)$,

$$\bullet \varphi_k(Q'_K, A_k, Q) = \|P_k - Q'_k A_k Q\|_{\varphi_{I\Lambda}}^2 = \|C'_k DC \Lambda^{-1} \Delta - Q'_k A_k Q \Delta\|_{\varphi_{II}}^2,$$

avec $\Delta^2 = \Lambda$.

$$\begin{aligned} \varphi_k(Q'_K, A_k, Q) &= \text{tr}(C'_k D [C \Lambda^{-1} C' D] C_k + Q'_k A_k Q \Lambda Q' A_k Q_k - 2C'_k DC Q' A_k Q_k) \\ &= \text{tr}(C'_k D (P_c C_k) + \Lambda Q' A_k^2 Q - 2C'_k DC Q' A_k Q_k) \end{aligned}$$

$P_c = C \Lambda^{-1} C' D$ définit le projecteur d'une variable sur le sous-espace F_s engendré par les variables du tableau C .

Si C est le tableau compromis et si $s = r$ (alors $s = q = r$), on a $c'_k \in F_s$, donc $P_c C_k = C_k$ et $\text{tr}(C'_k D P_c C_k) = \text{tr}(\Lambda_k) = \sum_{i=1}^{q_k} \lambda_i^k = \mathcal{I}_k$

Sinon $\text{tr}(C'_k D P_c C_k) \leq \mathcal{I}_k^s = \sum_{i=1}^s \lambda_i^k$, qui est l'inertie du nuage k pour $C_k (n \times s)$.

$$\bullet \Phi_k(Y_k, A_k, Y) = \|Y_k - Y A_k\|_{\varphi_{DI}}^2$$

$$\begin{aligned} \Phi_k(Y_k, A_k, Y) &= \text{tr}(C_k Q'_k Q_k C'_k D + C Q' A_k^2 Q C' D - 2Q_k C'_k DC Q' A_k) \\ &= \text{tr}(C'_k DC_k + \Lambda Q' A_k^2 Q - 2C'_k DC Q' A_k Q_k) \end{aligned}$$

$$\Phi_k(Y_k, A_k, Y) = \varphi_k(Q'_K, A_k, Q) + d_k, \text{ en posant } \mathcal{I}_k^s - \text{tr}(C'_k D P_c C_k) = d_k \geq 0.$$

En sommant selon k , on obtient (1) avec $\delta_s = \sum_{k=1}^m d_k \geq 0$. On notera $\Phi_s = \varphi_s + \delta_s$, où l'indice s rappelle que les calculs sont faits avec des tableaux C et C_k de dimensions $(n \times s)$.

2) C étant le tableau compromis, $s = q = r$ entraîne $d_k = 0$ et $\delta_s = 0$. On retrouve que l'égalité $P_k = Q'_k A_k Q$ est bien une propriété caractéristique des tableaux à composantes proportionnelles. ■

Cette remarquable propriété permet surtout de passer d'une caractérisation des tableaux à composantes proportionnelles, à la détermination des composantes

conjointes de l'AFJ. Il suffit d'emboîter deux algorithmes simples. Le premier, dit interne (exposé en **c1**, **c2**, **c3**), travaille sur des tableaux $(s \times s)$ pour déterminer (Q_k, A_k) , $k \in K$. Le second, dit externe (exposé en **c4**), traite des tableaux $(n \times s)$ pour déterminer Y et donc C .

c) L'algorithme de l'AFJ.

c1) Principe et initialisation de l'algorithme interne.

Les tableaux C et C_k , $k \in K$ étant donnés, l'algorithme est fondé sur la possibilité de modifier un tableau Q_k , A_k ou Q de façon à faire décroître la fonction φ_s et donc aussi Φ_s . De même, une modification du tableau C pourra faire décroître Φ_s .

Comme le note Ten Berge (1986, 1988), cela montre qu'il existe des solutions, puisqu'on construit une suite de valeurs décroissantes pour Φ_s qui est minorée par 0, et fournit une solution dans le même temps. Celle-ci peut être un minimum local ou global, d'où l'idée de faire plusieurs initialisations différentes pour comparer les résultats.

Pour des composantes quasi-proportionnelles l'initialisation de C par le tableau compromis de l'AFJ1 nous met suffisamment près de l'optimum global pour éviter les minima locaux liés à l'algorithme externe.

On peut alors fixer presque arbitrairement $Q_k, A_k, k \in K$ pour chercher Q . Il semble que le choix $Q_k = I_s$ et $A_k = \text{diag}(P_k)$, $k \in K$, avec renormalisation des A_k telle que $\sum_k A_k^2 = mI$, soit simple et efficace. Dans le cas où une composante c_k^j serait proportionnelle à c^i (avec $j \neq i$) on pourrait trouver une initialisation meilleure, mais plus complexe. Cela se remarque à la forme particulière du tableau P_k . On peut aussi envisager plusieurs initialisations de l'algorithme interne pour vérifier qu'on atteint bien un minimum global.

c2) Calcul de Q minimisant φ_s .

Pour (Q_k, A_k) fixés, $k \in K$ avec $\sum A_k^2 = mI$. On a :

$$\Phi_s = \text{tr}\left(\sum_{k=1}^m \Lambda_k^2 + m\Lambda\right) - 2 \text{tr}\left(\sum_{k=1}^m C_k' DC Q' A_k Q_k\right) = \varphi_s + \delta_s$$

et avec $C_k' DC = P_k \Lambda$ on a :

$$\text{Min}_Q \varphi_s \iff \text{Max}_Q \text{tr}\left(Q' \sum_k (A_k Q_k P_k \Lambda)\right)$$

Avec la décomposition en valeurs singulières on écrit : $\sum_k (A_k Q_k P_k) \Lambda = V \Delta U'$. En appliquant le théorème de Kristof (Kristof 1970, Le Calvé 1976, Ten Berge 1983), on trouve que l'extremum recherché s'obtient avec $Q = VU'$.

c3) Calcul de A_k et Q_k minimisant φ_s .

On détermine A_ℓ , puis Q_ℓ , pour tous les éléments successifs ℓ de K en supposant fixé $Q, (A_k, Q_k), k \in K - \{\ell\}$

- Min $\tau(A_\ell) = \text{Min}[\text{tr}(A_\ell^2 Q \Lambda Q') - 2 \text{tr}(A_\ell Q_\ell P_\ell \Lambda Q')]$ (Q_ℓ fixé). Ce minimum s'obtient en annulant toutes les dérivées :

$$\frac{\partial \tau}{\partial a_{\ell j}} = 2a_{\ell j} (Q \Lambda Q')_j^j - 2(Q_\ell P_\ell \Lambda Q')_j^j, j \in [s]$$

d'où :

$$A_\ell = \text{diag}(Q_\ell P_\ell \Lambda Q') [\text{diag}(Q \Lambda Q')]^{-1}, \ell \in K$$

- Min $\nu(Q_\ell) \iff \text{Max tr}(Q_\ell P_\ell \Lambda Q' A_\ell)$ (A_ℓ fixé). Avec la *s.v.d.* on a : $P_\ell \Lambda Q' A_\ell = V_\ell \Delta_\ell U'_\ell$. On obtient :

$$Q'_\ell = V_\ell U'_\ell \quad \text{ou} \quad Q_\ell = U_\ell V'_\ell, \ell \in K$$

On réitère le calcul de Q et $(A_k, Q_k), k \in K$ tant que φ_s décroît de façon significative. La fonction $\Phi_s = \sum_k \|Y_k - Y A_k\|^2$ a évidemment subi la même décroissance et on la calcule uniquement lors de cette dernière étape.

c4) L'algorithme externe : calcul de C .

Pour Y_k donc Q_k et A_k fixés, $k \in K$, on calcule une nouvelle valeur de Y faisant décroître Φ_s . C'est : $Y = \frac{1}{m} \sum_k Y_k A_k$ en vertu de la proposition 10.

On en déduit, au moyen de la *s.v.d.* par exemple, un nouveau tableau $C(n, s)$ qui contient les composantes principales du triplet (Y, I, D) . La *s.v.d.* de Y s'écrit $Y = \Gamma \Delta U'$. On prendra $C = \Gamma \Delta$ et la matrice Q telle que $Y = C Q'$ est évidemment $Q = U$.

Avec le nouveau tableau C , on réitère l'algorithme interne pour s'arrêter lorsque la fonction $\Phi_s = \sum_k \|Y_k - Y A_k\|^2$ ne décroît plus de façon significative. Notons qu'après cette première itération les tableaux $P_k = C'_k D C \Lambda^{-1}$ ne représentent plus les \sqrt{m} -projections $\sqrt{m} \tilde{U}'_k M U$ puisque C et U ne sont plus relatifs au nuage compromis de l'I.C.

3.3.- Les résultats de l'AFJ.

a) Interprétation de la méthode

- A l'optimum on connaît $Y, C, Q, (Q'_k, A_k, C_k), k \in K$.

On connaît donc l'I.E. : $\hat{Y}_k = Y_k = C_k Q'_k, k \in K$ et $\bar{Y} = \frac{1}{m} \sum_k Y_k$;

on connaît aussi l'I.D. : $\hat{X}_k = Y_k A_k = C_k Q'_k A_k, k \in K$ et $\hat{X} = \frac{1}{m} \sum_k \hat{X}_k = Y$.

- Si les tableaux sont à composantes proportionnelles, l'I.D. est confondue avec l'I.C. et le tableau compromis C vérifie $C = YQ$. L'AFJ se résoud sans itérer l'algorithme externe. Cela nous montre que la méthode est fondée sur la recherche de l'**Intrastructure Déformante** (dont le tableau Y) et le fait que celle-ci détermine immédiatement l'**Intrastructure Equisecteur**.

- Si les tableaux sont à composantes quasi-proportionnelles, on montre une très grande proximité entre les sous-espaces E_C et E_D , avec des particularités décrites en Annexe IV. C'est pourquoi l'initialisation de l'algorithme avec le tableau compromis nous met très près de l'optimum. L'algorithme externe se réitère d'autant moins que les composantes sont proches de la proportionalité.

- Par analogie à l'I.C. ou l'I.D.P. (cf Définition 8) nous présentons l'I.E. dans la base de E_E portant les axes principaux du nuage moyen, nommée Repère Equisecteur Principal (R.E.P.). L'expression de l'I.E. Principale est alors : $\hat{C}_k = Y_k Q_E = C_k Q'_k Q_E, k \in K$ et $\bar{C} = \bar{Y} Q_E$. Dans cette I.E.P. nous indiquons les directions des axes Δv_j qui définissent les composantes conjointes $Y_k = Y A_k + \varepsilon_k$ avec $\sum_k \|\varepsilon_k\|^2$ minimum.

b) Applications

Toutes les données traitées par INDSCAL, Statis ou l'AFM, en particulier les tableaux quantitatifs $X_k(n \times p_k)$ ou $(n \times p)$ auront avantage à être traités par l'AFJ.

- Le critère minimisé par l'AFJ est voisin de celui de INDSCAL qui s'écrit : $\text{Min} \sum_k \|W_k - \hat{W}_k\|^2$. Nous obtenons au plan théorique la démonstration de l'existence de l'optimum et un algorithme simple à convergence très rapide. Au plan pratique nous reconstituons $\hat{W}_k = Y A_k^2 Y'$, $k \in K$ avec en plus une **Intrastructure Exacte** : $\hat{Y}_k = Y_k = Y A_k + \varepsilon_k$ et la position des axes de plus grande proportionalité relativement à ceux de plus grande inertie.

• Nous avons vu que les infrastructures de Statis ou de l'AFM étaient inexactes. Nous y avons ajouté la possibilité d'en déduire une infrastructure exacte avec une formulation précise de la géométrie le permettant, et les conditions nécessaires à l'obtention de ce résultat. Elles n'ont presque rien à voir avec le choix des poids ρ_k affectés aux tableaux. A priori, les meilleurs poids sont ceux qui donnent aux tableaux $\rho_k X_k$ la même inertie mais, d'un autre côté, les poids égaux sont les plus simples. Pour des données non-cubiques, les tableaux sont centrés et réduits et une prémultiplication par $1/p_k$ assure les deux conditions. Avec les données cubiques, pour éviter la prémultiplication des tableaux on peut prendre des poids proportionnels à $1/\|X_k\|$ (cf Chap V).

• L'interstructure de Statis offre un critère suffisant pour que l'Infrastructure soit significative : il suffit de vérifier que le premier axe de l'interstructure (nuage des W_k) porte plus de 70% de l'inertie totale. La dispersion des W_k n'est donc pas gênante à condition qu'elle se situe essentiellement dans une seule direction. Ce point sera développé au chapitre V.

c) L'Analyse Procuste.

Avec $T(\sum_k q_k \times q_m)$, $q_m = \text{Inf}_k q_k$, $T' = (T'_1 \quad \dots \quad T'_m)$ et $T'_k T_k = I_{q_m}$ le problème procuste asymétrique s'écrit aussi $\text{Min } \psi(T) = \sum_k \sum_\ell \|C_k T_k - C_\ell T_\ell\|^2$. L'interstructure associée n'est pas exacte et on peut maintenant lui substituer deux problèmes procustes symétriques. Le premier déjà signalé correspond à $T(mq \times q)$ avec $q = \sup_k q_k$ et $T'_k T_k = I_q$. Il porte sur les tableaux $C_k = (C_k \quad 0)$ ($n \times q$), $k \in K$. On peut aussi considérer les tableaux $C_k(n \times s)$ limités aux s premières composantes, à condition qu'ils contiennent une grande part de l'inertie des nuages N_k (plus de 95%), et faire une analyse procuste symétrique qui perd très peu d'information. On aura alors $T(ms \times s)$. Dans la suite nous ne distinguerons plus ces deux cas.

Proposition 12.- *La minimisation de la fonction $\psi(T)$ de l'Analyse Procuste Symétrique (A.P.S.) équivaut à la minimisation de la fonction : $\Phi(Y_k, I, Y) = \sum_k \|Y_k - YI\|_{\varphi_{D_i}}^2$, qui est la fonction de l'AFJ contrainte par $A_k = I$, $k \in K$. L'APS est donc une AFJ contrainte.*

■ En introduisant des poids ρ_k égaux on montre que :

$$\sum_k \sum_\ell \|C_k T_k - C_\ell T_\ell\|^2 = 2m \sum_k \|C_k T_k\|^2 - \frac{1}{m} \sum_\ell \|C_\ell T_\ell\|^2$$

Dans l'AFJ, on a : $Y_k = C_k Q'_k = C_k T_k$ avec $T_k = Q'_k$, $k \in K$ et $Y = \frac{1}{m} \sum_k Y_k A_k$.
 Si $\forall k \in K A_k = I$, on a : $Y = \frac{1}{m} \sum_k C_k T_k$. Et on peut écrire :
 $\Phi(Y_k, I, Y) = \sum_k \|Y_k - Y I\|_{\varphi_{DI}}^2 = \sum_k \|Y_k Q - Y Q\|_{\varphi_{DI}}^2 = \sum_k \|Z_k - C\|_{\varphi_{DI}}^2$. ■

Définition 11.- De même que les composantes conjointes $Y_k = Y A_k + \varepsilon_k$ sont les composantes les plus proches de la proportionalité, on appelle **Composantes Procustes** d'une famille de triplets (X_k, M_k, D) , équivalents à (Z_k, I, D) les composantes des tableaux $Z_k = C + \varepsilon_k$ minimisant $\sum_k \|\varepsilon_k\|^2$ avec $C = \frac{1}{m} \sum_k Z_k$.
 Ce sont les composantes les plus proches, au sens des moindres carrés, de l'égalité à des composantes principales.

L'APS, en tant qu'AFJ contrainte par $A_k = I$, vérifie les mêmes propriétés : elle définit une infrastructure exacte par \sqrt{m} -projection sur le sous-espace Procuste E_P . L'introduction de la contrainte $A_k = I$ dans l'algorithme conduit à la solution, mais moins rapidement que dans l'AFJ classique car E_C est "moins proche" de E_P que de E_D . L'Infrastructure Procuste est moins significative que l'I.E. car bien évidemment $\sum_k \|Y_k - Y A_k\|^2 \leq \sum_k \|Z_k - C\|^2$. L'AFJ devrait donc être préférée à l'Analyse Procuste, et à défaut, l'algorithme de l'AFJ contrainte par $A_k = I$ donnera sans doute d'aussi bons résultats que les algorithmes classiques d'A.P. du fait qu'il travaille souvent sur des tableaux $(s \times s)$ et non $(n \times s)$. La recherche pourra être approfondie dans ce sens.

Pour conclure cette partie théorique, rappelons que l'AFJ donne une infrastructure exacte et la plus significative possible puisqu'on superpose dans l'I.E. les axes associés aux variables les plus proches de la proportionalité. Cette infrastructure est parfaitement significative lorsque le critère e vaut 0 (composantes proportionnelles) et elle l'est d'autant moins que e s'accroît avec le risque de voir apparaître une multiplicité d'infrastructures également significatives. Nous pensons qu'une utilisation intensive de la méthode permettra d'éclairer ce problème particulier. Dans les trois applications suivantes, l'AFJ donne d'excellents résultats qui soulignent bien les limites des méthodes antérieures.

CHAPITRE IV :

LES INTRASTRUCTURES SUR TROIS EXEMPLES

Dans la présentation suivante des exemples, l'objectif habituel de mise en évidence des phénomènes essentiels contenus dans les données n'apparaîtra qu'au second plan. Notre objectif est d'exposer les différents types de relations entre les Intrastructures Compromis et Equisecteur. De ce point de vue il est inutile de discuter du choix des métriques (qui peuvent transformer l'ACP en AFC), ou du traitement préalable des données : effectifs bruts ou profils. Nous restons dans l'ACP classique (métrique I) et pour l'exemple des cantons, nous traitons le cas des profils et des effectifs bruts. Les tableaux sont évidemment centrés et nous avons choisi pour l'AFJ de ne pas réduire les variables (ce qui s'imposerait pour les données non-cubiques) afin de traiter strictement les mêmes données par les trois méthodes présentées. Nous avons donc renoncé à la normalisation des données permettant d'aborder l'AFJ dans les meilleures conditions. Cela ne nous empêche pas d'obtenir d'excellents résultats.

Nous développons successivement les trois exemples en présentant d'abord les données (cubiques) et les résultats de l'ACP du Tableau Conjoint Empilé. Puis viennent les résultats de l'AFJ qui permettent d'expliquer ceux qu'on obtient par l'AFJ1, en particulier les déformations de l'I.C..

4.1.- Les accidents

a) Des données cubiques

Le terme accidents remplace pudiquement le mot tués. En effet, nous disposons ici d'une famille de 9 tableaux donnant pour les années 1978 à 1986 et pour 22 régions françaises (les individus), les 9 variables nombre de tués sur la route selon le type d'usager. Ces données cubiques sont présentées ci-contre (tableau Tab. 1) d'une façon inhabituelle puisqu'on donne successivement pour chaque région (codes 02 à 23) les valeurs des 9 variables pour les années 1978 à 1986 (codes 8 à 6).

Nous précisons ci-dessous les codes des variables et des régions :

Variables	Régions	
<i>on</i> piétons	02 Nord-Pas de Calais	13 Midi-Pyrénées
<i>bi</i> bicyclettes	03 Picardie	14 Champagne-Ardennes
<i>cy</i> cyclomoteurs	04 Ile de France	15 Lorraine
<i>mo</i> motos	05 Centre	16 Alsace
<i>vl</i> véhicules légers	06 Haute Normandie	17 Franche-Comté
<i>ca</i> camionnettes	07 Basse Normandie	18 Bourgogne
<i>pl</i> poids lourds	08 Bretagne	19 Auvergne
<i>tc</i> transport en commun	09 Pays de Loire	20 Rhône-Alpes
<i>au</i> autres	10 Poitou-Charentes	21 Languedoc-Roussillon
	11 Limousin	22 Provence-Côte d'Azur
	12 Aquitaine	23 Corse

Le tableau ci-dessous des totaux des colonnes pour chacune des 9 années montre l'évolution globale suivante :

Baisse (plus ou moins régulière) pour *on*, *bi*, *cy*, *ca* et *tc*.

Hausse pour *vl* (environ 10 %) et baisse du total de 9 %, donc une hausse relative importante de la catégorie *vl*.

année	<i>on</i>	<i>bi</i>	<i>cy</i>	<i>mo</i>	<i>vl</i>	<i>ca</i>	<i>pl</i>	<i>tc</i>	<i>au</i>	total
1	2125.	571.	1618.	782.	6233.	260.	214.	35.	119.	11957.
2	2092.	589.	1381.	970.	6546.	265.	202.	38.	116.	12197.
3	2196.	659.	1271.	1059.	6699.	267.	208.	34.	113.	12510.
4	1945.	555.	1362.	804.	6941.	234.	215.	25.	171.	11953.
5	1910.	501.	1107.	813.	7109.	214.	183.	64.	129.	12030.
6	1879.	536.	955.	846.	6956.	152.	198.	27.	128.	11677.
7	1760.	465.	864.	820.	7121.	193.	159.	18.	125.	11525.
8	1557.	436.	797.	810.	6419.	161.	148.	15.	114.	10447.
9	1639.	438.	714.	790.	6867.	170.	207.	14.	122.	10961.

b) L'Intrastructure de l'ACP (I.ACP) du Tableau Conjoint Empilé

En traitant les effectifs bruts, nous avons un effet de taille très marqué permettant de représenter 99.2 % de l'inertie totale sur les trois premiers axes. Ainsi l'intrastructure (figure 1) donne une description très précise de l'évolution

des diverses catégories de tués (sauf celles ayant de petits effectifs) pour chacune des régions.

Chaque trajectoire joint les points correspondant aux années successives, l'origine (1978) étant repérée par la mention du numéro de région. Certaines trajectoires sont ici illisibles du fait des impératifs de la reproduction : réduction des formats et disparition (éventuelle) des couleurs. Nous ne parlerons ici que des trajectoires lisibles sachant qu'une étude détaillée a été menée à l'aide de dessins en grand format.

L'orientation des plans principaux est facilitée par la projection des 9 vecteurs unitaires (appelés individus-variables en Annexe I) portant les axes de l'espace des individus. Ceux qui sont bien représentés dans un plan sont au bord d'un cercle de rayon unité. On a les valeurs :

Axes	% inertie	<i>on</i>	<i>bi</i>	<i>cy</i>	<i>mo</i>	<i>vl</i>	<i>ca</i>	<i>pl</i>	<i>tc</i>	<i>au</i>
1	93.26	.300	.067	.136	.153	.928	.028	.024	.003	.011
2	98.24	-.876	-.070	-.166	-.268	.357	.009	.020	.009	.001
3	99.22	.102	-.330	-.860	.361	.058	-.020	.010	.020	-.055

Les directions des variables de forte variance sont les mieux représentées en particulier *vl* (avec $\|\hat{vl}\| = 0.996$ dans le plan 1-2) puis *on* pour le plan 1-2 et *vl*, *cy*, *mo* dans le plan 1-3. Les autres individus-variables dont la projection est presque confondues avec l'origine, ne sont pas représentées.

Les axes principaux ont ainsi les significations suivantes :

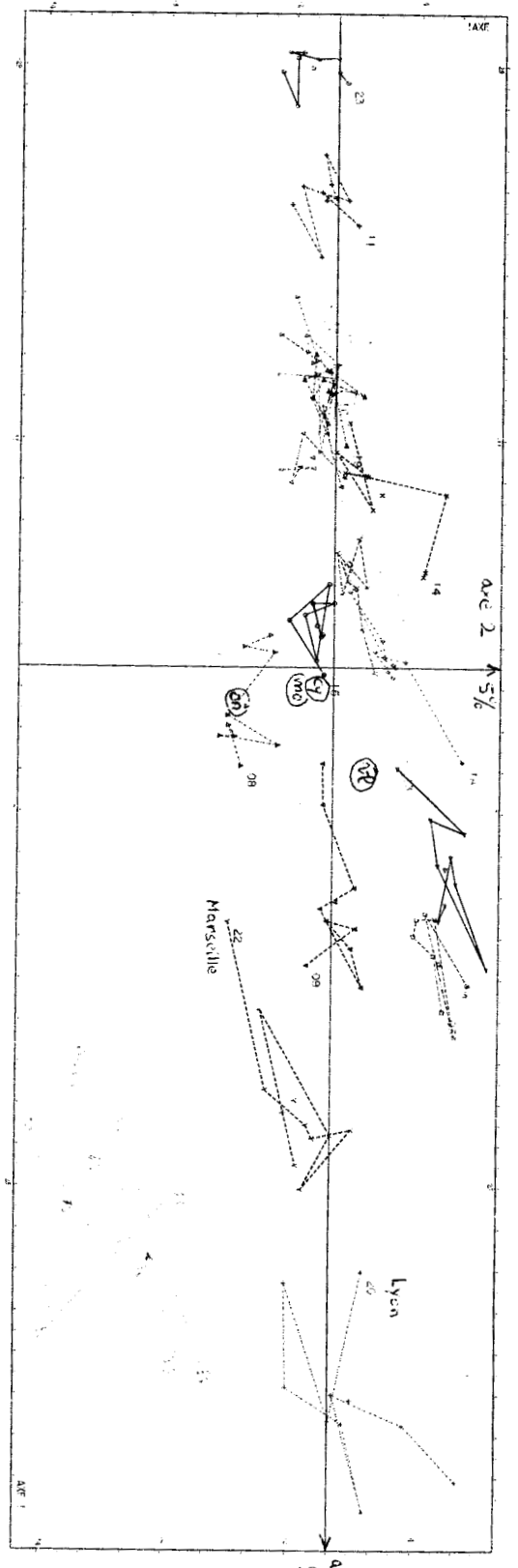
- L'axe 1 est un axe de taille, facile à interpréter : il oppose les régions avec beaucoup de tués à celles où il y en a peu.
- L'axe 2 oppose les régions où domine *vl* (en valeur relative) aux régions pour lesquelles la proportion des catégories *on*, *bi*, *cy* ou *mo* est plus forte.
- L'axe 3 distingue alors les dominantes *cy* (ou *bi*) de la dominante *mo*.

La lecture de cette instructure nous renseigne immédiatement sur l'essentiel des phénomènes contenus dans ces données.

Pour chaque région on voit ses caractéristiques moyennes en regardant la position du compromis et le sens de son évolution en suivant la trajectoire. Ainsi les régions 02 (Nord/Pas-de-Calais) et 08 (Bretagne) ont un nombre total de tués

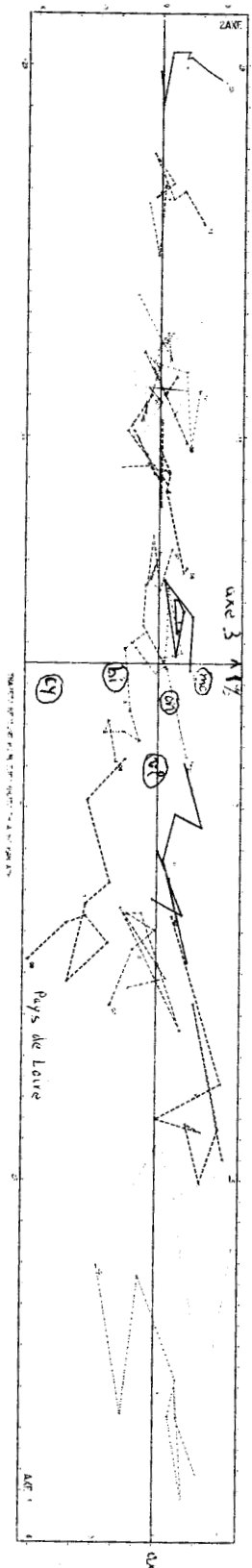
Figure 1

1.ACP pour les accidents



FRANCE REPUBLIQUE FRANÇAISE

effectifs divisés par 10; 22 régions



situé dans la moyenne, avec une forte proportion des catégories *on*, *mo*, *cy* (plan 1-2) et surtout *cy* ou *bi* (plan 1-3). L'évolution de ces régions montre une diminution du total (relativement à l'ensemble) et un léger rééquilibrage des diverses catégories. Pour 04 (Ile-de-France) il y a un nombre important de tués avec une importance exceptionnelle des catégories *on*, *mo*, *cy* (plan 1-2) et plus particulièrement *on*, *mo* (plan 1-3). On peut aussi détailler l'évolution année par année :

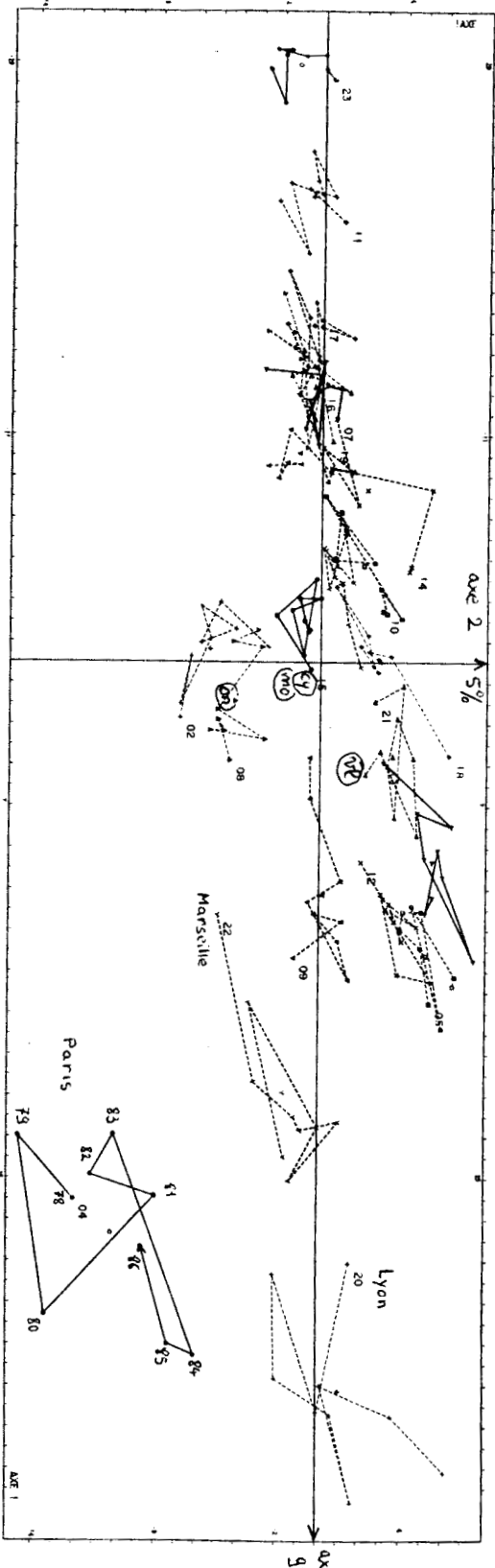
- 78-79, plan (1-2) : forte diminution de *vl*, léger accroissement des autres.
plan (1-3) : cet accroissement concerne *on* et *mo*.
La lecture des tableaux de données permet de vérifier cette évolution :
vl : 481 \rightarrow 458 ; *on* : 277 \rightarrow 290 ; *mo* : 106 \rightarrow 127 .
- 79-80, plan (1-2) : énorme accroissement pour *vl*.
plan (1-3) : accroissement de *on* et *mo* (éventuellement).
vérification : *vl* : 458 \rightarrow 552 (+20 %) ; *on* : 290 \rightarrow 318 ; *mo* : 127 \rightarrow 129.
- etc...

Rappelons que l'évolution d'une région, d'une année à la suivante, est définie relativement au centre de gravité du nuage des régions puisque ceux-ci sont tous positionnés à l'origine. L'évolution des centres de gravité g_k (cf. DACP Bourouche 1975) est ici modeste : il y a en moyenne 4.8 tués de moins par région et par an. Mais ce nombre est important pour les régions à petits effectifs, comme 23 (Corse) où le total des tués oscille entre 50 et 70 par an. Dans ce cas, l'évolution repérable sur la trajectoire qui est définie par la suite des points $(x_{23}^k - g_k)$, $k \in K$, reflète l'évolution de la Corse combinée à l'opposé de l'évolution des centres de gravité éventuellement plus importante. Ce constat peut inciter à introduire un point fixe dans les tableaux non centrés ($x_{n+1}^k =$ vecteur cst) en individu supplémentaire. Sa trajectoire sera celle de $(-g_k)$, $k \in K$, à une translation près.

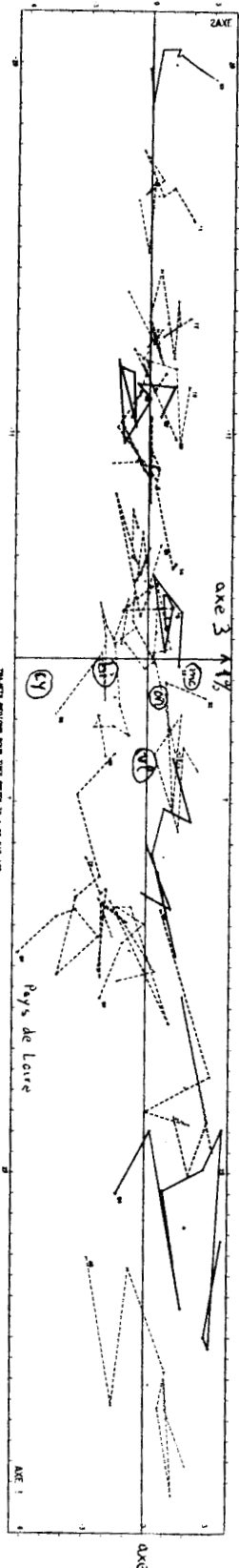
c) L'Infrastructure Equisecteur : des composantes principales proportionnelles

Pour appliquer l'AFJ aux données précédentes, considérées comme non-cubiques, on doit vérifier que la méthode est applicable. L'ACP du tableau moyen X_p donne pour les axes successifs, les cumuls d'inertie suivants : 92%, 96.6 %, 97.7 %, 98.4 %, 98.7 %... Soit s le nombre de composantes principales retenues dans les tableaux C et C_k . Au lieu de prendre $s = \sup(pk) = q = 9$ on peut accepter ici

I.ACP pour les accidents



effectifs divisés par 10; 22 régions



$s = 3$, puisque par ailleurs, au niveau de chaque nuage, les 3 premiers axes portent les parts d'inertie suivantes : 99.24 %, 99.12 %, 99.28 %, 99.57 %, 99.21 %, 99.43 %, 99.53 %, 99.30 %, 99.62 %. On a bien p_s et p_s^k supérieurs à $p_s^0 = 95\%$ le seuil jugé raisonnable. Rappelons aussi que les 3 premiers axes de l'I.E. expliquent 99.37 % de l'inertie de la réunion des nuages (moyenne des parts d'inertie précédentes), donc un peu plus que l'I.ACP donnant 99.22 %.

Dans cet exemple ce sont les composantes principales qui vérifient la proportionalité, avec une très petite erreur comme on va le voir avec les résultats de l'AFJ (Tab. 2).

La proportionalité des premières composantes principales se constate dès le départ de l'AFJ en regardant les tableaux $P_k = C'_k D C \Lambda^{-1}$, $k \in K$. L'égalité $P_k(1, 1) = a_{k1}$, pour tout $k \in K$, implique $\frac{1}{m} \sum P_k^2(1, 1) = 1$ qui est caractéristique de cette proportionalité. On a ici : $\frac{1}{m} \sum P_k^2(1, 1) = 1.000$; $\frac{1}{m} \sum P_k^2(2, 2) = 0.976$; $\frac{1}{m} \sum P_k^2(3, 3) = 0.717$, donc une proportionalité quasi parfaite des premières composantes principales, bonne des secondes, approximative des troisièmes.

Cette proportionalité se retrouve dans l'expression des matrices Q_k telles que $Y_k = C_k Q'_k$ et Q (t.q $Y = C Q'$) qui sont éditées avec les lignes mises bout à bout. Les Q_k et Q sont proches de l'identité car elles ont toutes des 1 sur la diagonale à 6.10^{-3} près. Excepté pour les tableaux de 78 ou 86, ce sont bien les composantes principales qui sont proportionnelles et l'erreur relative moyenne : $e = \sum_k \|Y_k - Y A_k\|^2 / m \|Y\|^2 = 2.3\%$ est minime.

On peut interpréter les coefficients de proportionalité a_{kj} en fonction de la signification donnée aux axes principaux qui sont ici les axes de proportionalité. Il suffit de rappeler que pour des composantes proportionnelles, les variances sont proportionnelles à a_{kj}^2 . Dans ces conditions :

- Les a_{k1} , $k \in K$, reflètent l'accroissement de la variance de la variable vl qui est très corrélée avec la première composante principale (cet accroissement résulte de l'accroissement de cette variable dans certaines régions comme 20-22).
- Les a_{k2} , $k \in K$, montrent que la dispersion le long de l'axe 2 tend à diminuer. L'opposition s'atténue entre régions à forte proportion de la variable vl , et régions où les variables on , bi , cy , mo connaissent une proportion supérieure à la moyenne nationale. Leur variance diminue et les profils de ces catégories tendent à se ressembler.

Tableau 2

Les tableaux $P_k (3 \times 3)$ avec $P_k = C_k' D C \Lambda^{-1} = \sqrt{m} U_k' M U_k$, $k \in K$

Les lignes des P_k sont mises bout à bout.

L'AFJ recherche Q_k', Q, A_k tels que $\sum_k ||P_k - Q_k' A_k Q||^2$ soit minimale

P 1	0.895	-0.159	-1.289	-0.000	1.043	0.017	0.010	-0.051	0.873
P 2	0.958	0.053	-0.539	-0.006	1.214	-0.144	0.003	0.018	0.838
P 3	0.982	-0.419	-0.138	0.019	1.033	0.077	-0.001	-0.010	0.666
P 4	1.009	0.019	-0.292	-0.004	0.844	0.105	0.002	0.001	0.654
P 5	1.020	0.215	-0.279	-0.017	1.017	-0.055	0.002	0.016	0.827
P 6	1.056	0.327	0.012	-0.015	0.984	0.029	-0.001	-0.000	0.734
P 7	1.039	-0.171	0.537	0.004	0.813	0.076	-0.005	-0.018	0.636
P 8	0.986	-0.042	0.790	-0.004	0.873	-0.019	-0.002	0.039	0.618
P 9	1.045	0.161	0.980	-0.009	0.893	-0.261	-0.004	0.062	0.529
P 0	0.999	-0.002	-0.024	-0.003	0.968	-0.019	0.001	0.006	0.708
seuils b,d		0.523	2.614						

A 0 1.000 0.976 0.717 : valeurs de $\frac{1}{m} \sum_k P_k^2 (j,j)$, $j \in [3]$

Matrices Q et Q_k' , $k \in K$ avec les lignes mises bout à bout.

Q	1.000	0.006	0.009	-0.006	1.000	-0.005	-0.009	0.005	1.000
Q' 1	0.999	-0.014	-0.040	0.017	0.997	0.073	0.038	-0.073	0.997
Q' 2	1.000	0.001	-0.018	-0.001	1.000	-0.013	0.018	0.013	1.000
Q' 3	0.999	-0.045	-0.008	0.046	0.998	0.033	0.007	-0.033	0.999
Q' 4	1.000	-0.000	-0.012	0.001	1.000	0.026	0.012	-0.026	1.000
Q' 5	1.000	0.019	-0.015	-0.019	1.000	-0.016	0.015	0.017	1.000
Q' 6	1.000	0.021	-0.007	-0.021	1.000	0.011	0.008	-0.011	1.000
Q' 7	1.000	-0.015	0.000	0.014	0.999	0.041	-0.001	-0.041	0.999
Q' 8	1.000	-0.003	-0.001	0.003	0.999	-0.050	0.002	0.050	0.999
Q' 9	1.000	0.009	0.000	-0.009	0.994	-0.110	-0.001	0.110	0.994

Les tableaux A_k , $k \in K$,
avec sur chaque ligne
les éléments a_{kj} , $j \in [3]$
de la diagonale de A_k

A 1	0.894	1.072	1.299
A 2	0.958	1.238	1.169
A 3	0.982	1.075	0.864
A 4	1.009	0.864	0.898
A 5	1.020	1.036	1.101
A 6	1.056	1.011	0.968
A 7	1.039	0.836	0.898
A 8	0.986	0.891	0.891
A 9	1.046	0.909	0.804
A 0	1.000	1.000	1.000

$$A_0 = \frac{1}{m} \sum_k A_k^2$$

Les matrices $Q_k' Q$ permettent de définir l'I.E.P. par: $\hat{C}_k = C_k Q_k' Q$.

Rappelons qu'on a pour l'I.D.P.: $\hat{\hat{C}}_k = C_k Q_k' A_k Q$ et pour l'I.C.: $\hat{\hat{X}}_k = C_k P_k$

Q'kQ 1	1.000	-0.009	-0.030	0.011	0.998	0.067	0.030	-0.068	0.997
Q'kQ 2	1.000	0.007	-0.009	-0.007	1.000	-0.019	0.009	0.019	1.000
Q'kQ 3	0.999	-0.040	0.001	0.040	0.999	0.029	-0.002	-0.028	1.000
Q'kQ 4	1.000	0.005	-0.003	-0.005	1.000	0.020	0.003	-0.020	1.000
Q'kQ 5	1.000	0.025	-0.006	-0.025	0.999	-0.022	0.006	0.022	1.000
Q'kQ 6	1.000	0.027	0.002	-0.027	1.000	0.005	-0.002	-0.005	1.000
Q'kQ 7	1.000	-0.009	0.009	0.008	0.999	0.036	-0.010	-0.035	0.999
Q'kQ 8	1.000	0.002	0.008	-0.002	0.998	-0.055	-0.008	0.055	0.998
Q'kQ 9	1.000	0.015	0.009	-0.014	0.993	-0.116	-0.011	0.114	0.997

- Les a_{k3} , $k \in K$, diminuent régulièrement, montrant que s'atténue l'opposition entre régions avec une forte proportion *bi* et *cy* (02, 08, 09, 12, 05) et les autres régions. Pour ces régions ces moyens de locomotion ont sans doute régressé et il y a eu une amélioration de la prévention.

Il est alors intéressant de comparer l'I.ACP et l'I.E. de l'AFJ dans laquelle on indique les projections des directions des variables selon une technique qui sera expliquée ci-dessous dans la présentation de l'I.C.. Pour la lisibilité on a limité l'I.E. (et l'I.C.) à 12 trajectoires. Comme l'ACP et l'AFJ n'ont pas donné la même orientation aux axes, nous avons disposé les dessins pour qu'ils soient comparables sans avoir à tourner cet ouvrage.

On voit très clairement que l'Intrastructure Equisecteur de l'AFJ (Fig. 2) est l'Intrastructure de l'ACP lorsqu'on en élimine les mouvements de rotation des nuages dans le plan. Ces rotations des nuages sont approximativement celles qui, dans les plans 1-2 ou 1-3 permettent au point 23_k (Corse à la date k) de venir coïncider avec le point 23 compromis, leur centre étant 0. L'élimination de ces rotations ne fait pas perdre d'information car elles expriment ici le phénomène déjà décrit avec les coefficients a_{k2} et a_{k3} .

d) L'Intrastructure Compromis et l'effet PLM

d1) La possibilité théorique de l'effet "PLM" est exposée en Annexe IV.

Rappelons que pour voir apparaître l'effet PLM, c'est à dire des déformations très spécifiques dans l'Intrastructure Compromis, il faut :

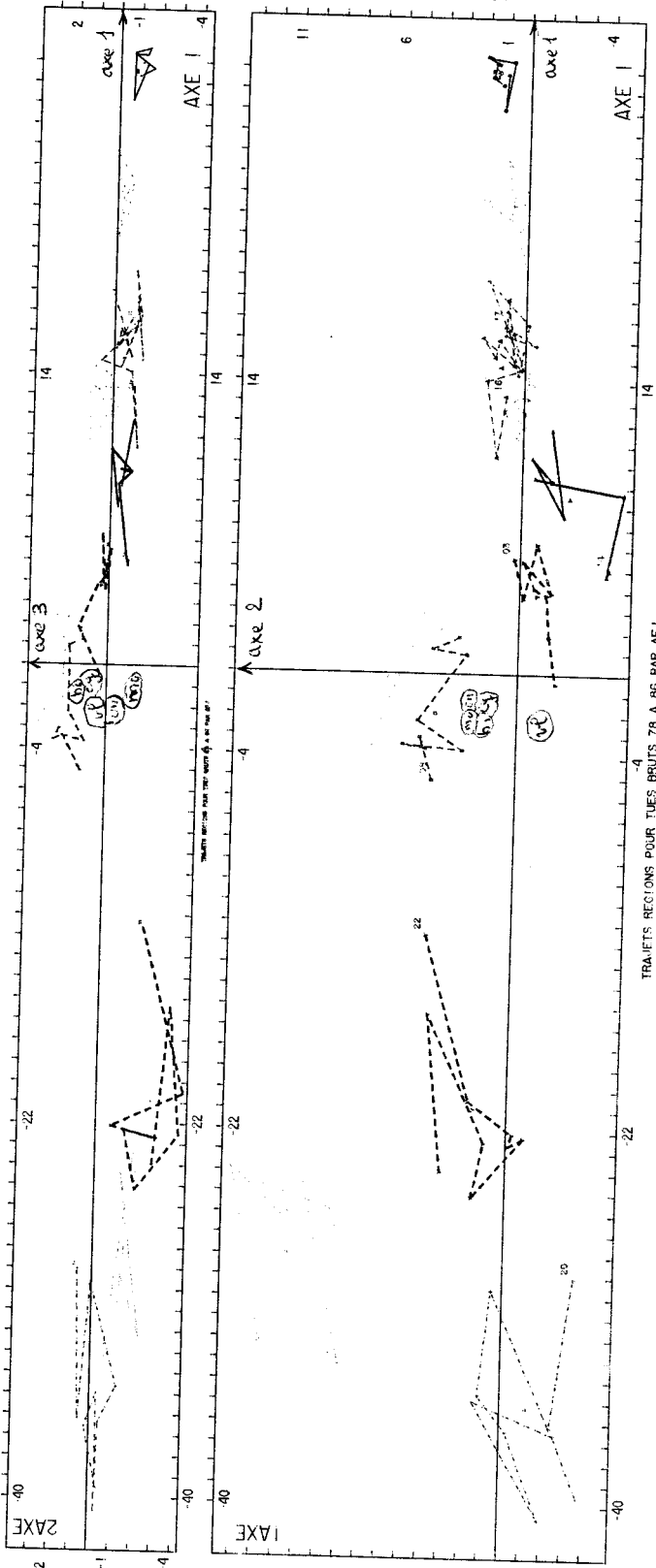
1) que le compromis d'une trajectoire ait une coordonnée particulièrement importante sur un axe Δv_j ($j > 1$) et contribue ainsi fortement à l'inertie portée par cet axe ;

2) que la coordonnée sur l'axe 1 des points de cette trajectoire vérifie l'égalité $y_i^{1k} = y_i^1 a_{k1} + \varepsilon_i^{1k}$ avec des erreurs (valeurs ε_i^{1k}) particulièrement importantes. Alors il y a déplacement du R.C. relativement au R.D., les axes Δv_j du R.C. se déplaçant vers les axes $\Delta \tilde{v}_j^k$ des nuages k pour lesquels ε_i^{1k} est important.

En regardant l'I.ACP (Fig. 1) on constate que deux trajectoires remplissent les conditions de l'effet PLM :

1) 04 ("Paris") pour le plan 1-2 est très excentrée et elle présente des évolutions sur l'axe 1 systématiquement opposées à celles de 20 ("Lyon") avec de très grandes

Figure 2 I.E. des accidents



effectifs divisés par 10; 12 régions représentées

oppositions en 80, 83 et 84, et aussi à celles de 22 ("Marseille"). Paris vérifie donc $y_4^{1k} = y_4^1 a_{k4} + \varepsilon_4^{1k}$ avec des ε_4^{1k} très importants.

Aussi l'axe Δv_2 du R.C. s'écarte de son homologue du R.D. pour s'approcher des axes $\Delta \tilde{v}_1^k$ des nuages N_k plongés dans E . On trouvera dans les composantes sur l'axe Δv_2 du R.C. une partie tenant à la composante y_i^{2k} du point considéré et une autre tenant sa composante y_i^{1k} , d'autant plus importante que celle-ci est grande. Cette deuxième partie n'existe évidemment pas sur l'axe Δv_2 du R.D. qui est dans un sous-espace orthogonal au sous-espace engendré par les axes $\Delta \tilde{v}_1^k$, $k \in K$. Ce phénomène se lit immédiatement dans l'intrastructure de l'AFJ1 (Fig. 3) avec l'amplification du phénomène quand y_i^{1k} s'accroît et qu'on s'éloigne de l'origine. Son nom "PLM" vient évidemment de "Paris Lyon Marseille".

2) 09 ("Pays de Loire") est excentrée sur l'axe 3 et connaît une forte évolution sur l'axe 1 s'opposant à presque toutes les régions. On obtient des déformations démesurées sur l'axe 3 de l'I.C., parfaitement proportionnelles aux coordonnées y_i^{1k} (Fig. 4).

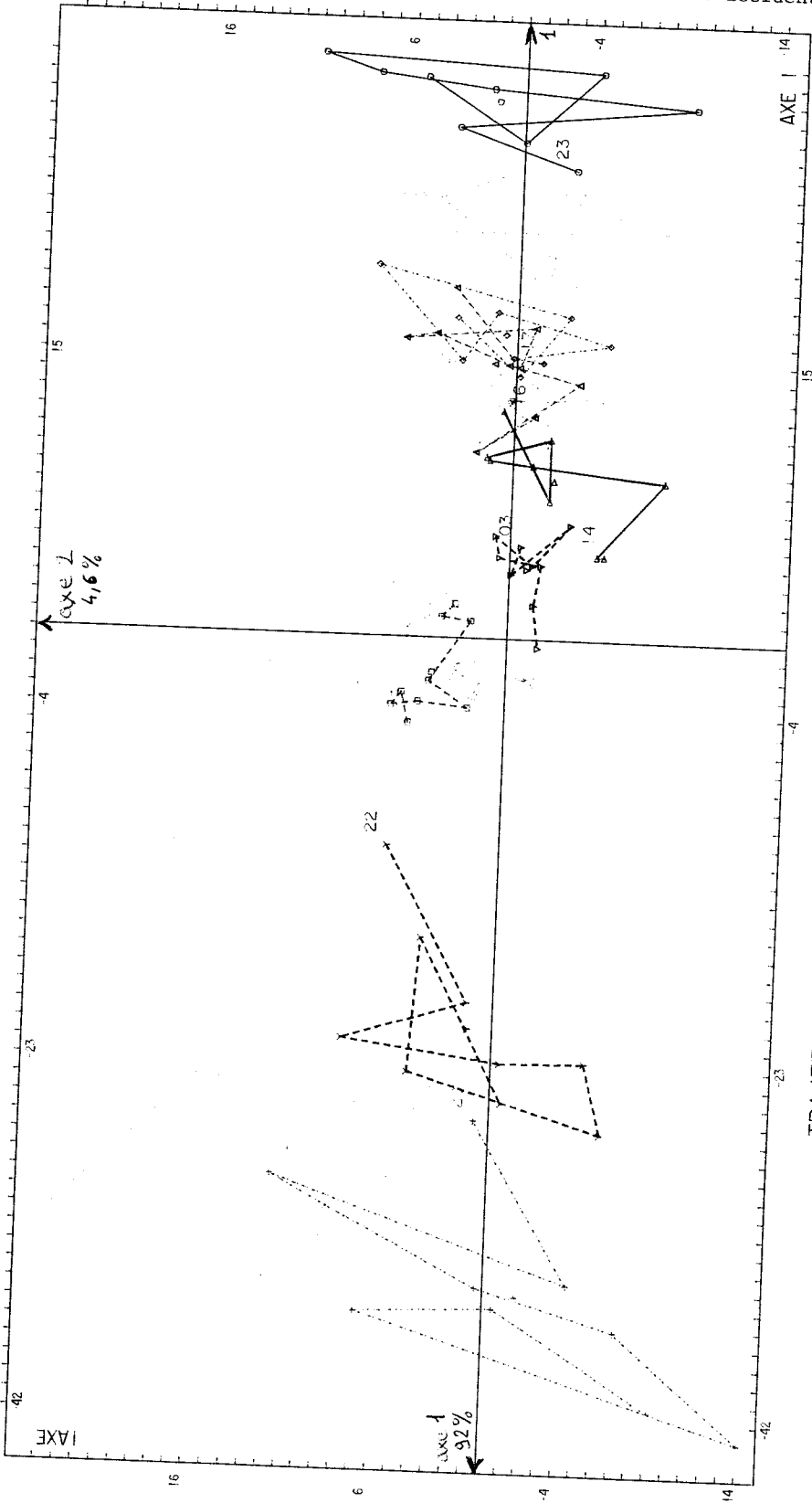
d2) Propriétés de l'I.C.

Sur cet exemple, rappelons quelques propriétés de cette Intrastructure Compromis de l'AFJ1 obtenue par la méthode Statis adaptée à la formulation donnée en 1.2.a.

Les pourcentages d'inertie portés par les 3 axes sont relatifs au seul nuage compromis positionné dans $E^p = \bigoplus^{\perp} E_k$. Ils sont légèrement inférieurs à ce que l'on obtient pour l'ensemble des nuages dans l'I.ACP ou l'I.E..

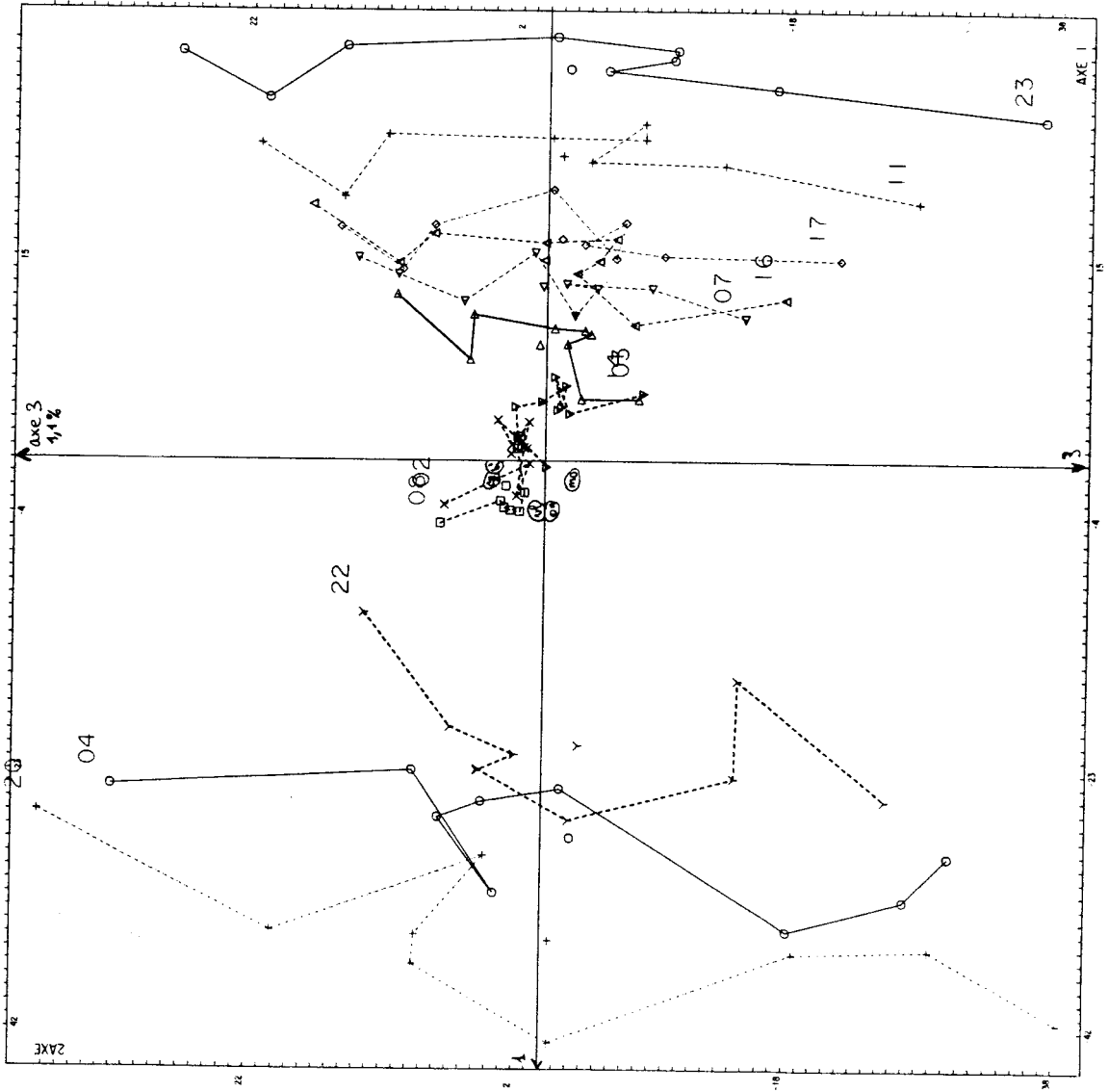
Le programme Statis2 donne également le cercle des corrélations pour l'ensemble des variables figurant dans le tableau X_ρ et donc pour des données cubiques, chaque variable est représentée m fois ($m = \text{card } K$). En considérant le barycentre des y_j^k , $k \in K$, on obtient une variable compromis y^j dans le cercle des corrélations dont on déduit (par la méthode exposée en annexe I) un axe Δe_j de l'espace des individus, lié par dualité à cette variable y^j , compromis des variables de même nom. Cette opération n'a de sens que pour les données cubiques et sous la condition que les variables de même nom soient suffisamment corrélées. Sinon on peut toujours représenter l'ensemble des directions des variables dans l'espace des individus.

Figure 3 I.C. des accidents



TRAJETS REGIONS POUR TUES BRUTS 78 A 80 PAR STATIS (AFJ1)
effectifs divisés par 10; 12 régions représentées

Figure 4 I.C. des accidents plan (1-3)



Pour les recensements de 1954, 1962, 1968 et 1975, on dispose de la répartition des effectifs de la population active en 9 catégories socio-professionnelles et 42 cantons ruraux du Languedoc-Roussillon. Les catégories socio-professionnelles étudiées sont :

<i>ea</i>	exploitants agricoles
<i>oa</i>	salariés agricoles
<i>ac</i>	patrons de l'industrie et du commerce
<i>pl</i>	professions libérales et cadres supérieurs
<i>cm</i>	cadres moyens
<i>em</i>	employés
<i>ou</i>	ouvriers
<i>se</i>	services
<i>cp</i>	autres catégories

Nous présentons encore (Tab. 3) les données en profils et dans les deux cas nous admettons qu'il s'agit de 4 tableaux de type individus (les cantons) \times variables quantitatives (les effectifs ou les profils). Ces tableaux doivent être centrés.

b) L'Intrastructure de l'ACP du TCE (cantons profils)

Dans les paragraphes 4.2.b. à 4.2.d., nous traitons des cantons en profils. Comme pour l'I.ACP précédente, nous donnons le pourcentage cumulé d'inertie expliquée par les axes, et le tableau des coordonnées des vecteurs Υ^j , projetés des e_j dans les plans principaux d'inertie.

Axes	inertie	<i>ea</i>	<i>oa</i>	<i>ac</i>	<i>pl</i>	<i>cm</i>	<i>em</i>	<i>ou</i>	<i>se</i>	<i>cp</i>
1	71.45	.677	.219	-.059	-.017	-.031	-.070	-.694	-.023	-.000
2	92.67	.564	-.753	-.048	-.014	.001	-.068	.325	-.020	.012
3	97.64	-.295	-.503	.473	.141	.231	.210	-.527	.179	.090

On a donc l'orientation suivante des axes principaux (Fig. 5) :

- l'axe 1 oppose les cantons ouvriers à gauche aux cantons agricoles à droite
- l'axe 2 oppose les ouvriers agricoles en bas aux ouvriers et exploitants agricoles en haut
- l'axe 3 oppose *ou*, *oa* et *ea* aux autres catégories.

Chaque trajectoire peut être regardée d'un double point de vue :

- La position de son centre de gravité dans chaque plan renseigne sur les particularités globales du canton. Ainsi D8 montre une très grande importance de la variable *ou* (proportion d'ouvriers).
- L'évolution au cours des 4 recensements qui indique (pour D8) :
 - une baisse de cette proportion (plan 1-2)
 - baisse confirmée dans le plan (1-3) qui laisse entrevoir une hausse de *oa*.

Le phénomène général dominant est ici que presque toutes les trajectoires convergent vers l'origine 0, ce qui indique une très nette tendance à homogénéiser les catégories existant dans les divers cantons. L'intérêt de l'intrastructure réside néanmoins dans sa capacité à détailler les évolutions individu par individu, ce que fait parfaitement cette IACP (à la petite taille du dessin près : Fig. 5).

c) L'Intrastructure Equisecteur et les calculs de l'AFJ

Les valeurs propres du compromis WD , associé au tableau X_p (avec $\rho_k = \frac{1}{m}$) donnent les parts d'inertie suivantes : 69.10 %, 88.83 %, 94.80 %, 97.20 %, 98.44 % ... Nous avons donc accepté de prendre $s = 3$ composantes (94.8 \approx 95 %). Mais il s'agit plutôt d'un choix pratique car nous ne sommes plus en mesure de garantir les inégalités $p_s^k > 95$ %, $k \in K$.

L'AFJ est encore très significative puisque l'erreur relative moyenne vaut 3.91 % avec des composantes conjointes qui ne sont absolument pas les composantes principales comme nous allons le voir. Nous présentons ici le détail des calculs de l'AFJ pour montrer la manière dont les deux algorithmes font converger vers la solution (voir Tab. 4).

L'algorithme interne, d'indice jc fait décroître le critère :

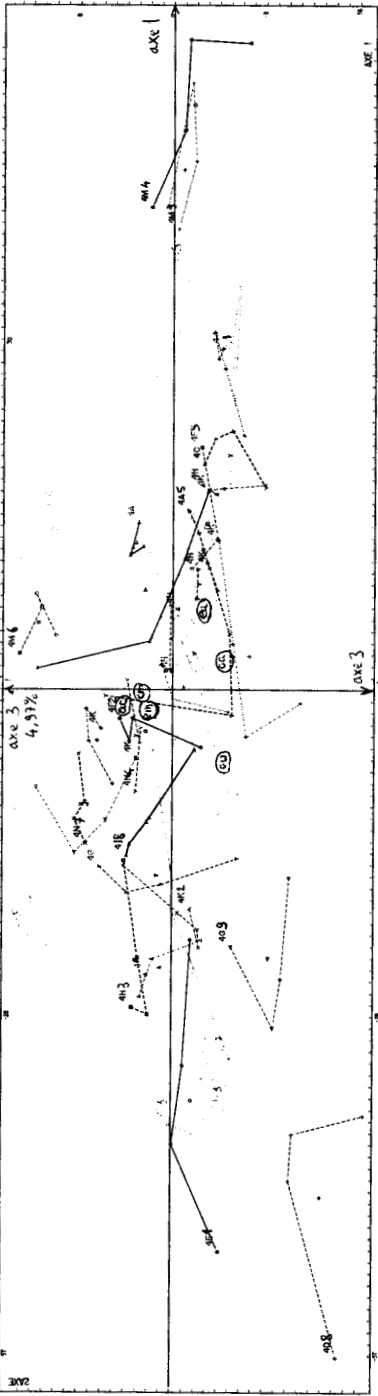
$$\varphi_s = \sum_k \|P_k - Q'_k A_k Q\|_{\varphi_{IA}}^2, \text{ et donc aussi le critère :}$$

$$\Phi_s = \varphi_s + \delta_s = \sum_k \|Y_k - Y A_k\|_{\varphi_{DI}}^2.$$

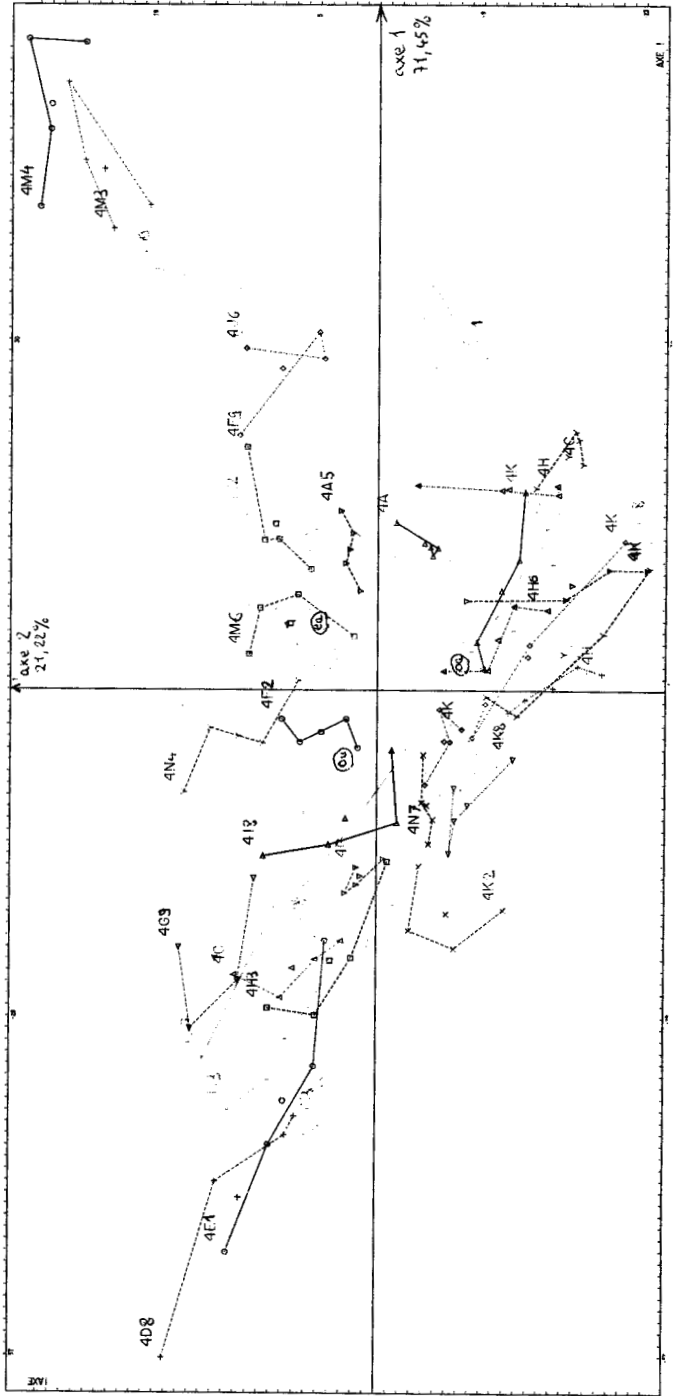
c1) Les calculs de l'AFJ

φ_s est donné après la première itération de l'algorithme interne. Lors de la dernière on donne successivement les valeurs pour Q , Q_k , A_k , $A_0 = \frac{1}{m} \sum_k A_k^2$, puis la valeur du critère φ_k de chaque tableau et φ_s , enfin, la valeur du critère Φ_k et Φ_s .

Figure 5 L'I.ACP des cantons en profils



TRAJECTOIRES CANTONS CSP PROFILS PAR ACP



L'algorithme externe, d'indice jm fait décroître le critère Φ_s en prenant $Y = \frac{1}{m} \sum_k Y_k A_k$ (Cf. 3.2.c4.) . Après la dernière itération ($jm = 3, jc = 2$), on retrouve les tableaux A_k normalisés, les erreurs relatives (par tableau et moyenne) les matrices $Q'_k Q$ puis pour chaque individu : les 3 premières composantes des 5 tableaux suivants :

$$\hat{X}_k = W_k D C \Lambda^{-1} = C_k P_k ; \hat{C}_k = C_k Q'_k Q ; C_k ; Y_k = C_k Q'_k ; Y A_k, k \in K.$$

et les composantes des deux compromis \hat{X}, \hat{C} ainsi que celles de Y . Ici, elles ne sont données que pour le premier individu A5.

Précisons que les seuils b et d relatifs aux algorithmes interne et externe sont fixés comme suit :

$d = 5.10^{-5} m \|C\|^2$, et on reprend l'algorithme externe tant que : $\Phi_{jm-1} - \Phi_{jm} > d$, avec $\Phi_{jm} = \sum_k \|Y_k - Y A_k\|^2$ à l'étape jm . Cela revient à continuer tant que la moyenne de l'écart relatif sur un élément $|(Y_k - Y A_k)_i^j / Y_i^j|$ diminue de plus de 0.7 %, en moyenne.

$b = d/5$ pour l'algorithme interne, car le critère : $\varphi_s = \sum_k \|P_k - Q'_k A_k Q\|_{\varphi_{j\Lambda}}^2$ est plus facile à faire décroître que le précédent. Il nécessite la s.v.d. de matrices $(s \times s)$ et non $(n \times s)$.

Ces seuils donnent de bons résultats pratiques, mais on peut être plus sévère sans autre inconvénient que d'allonger le temps de calcul. Les deux résultats essentiels pour interpréter l'I.E. de l'AFJ sont les tableaux Q et $Q_k, k \in K$, puis les tableaux $A_k, k \in K$.

Les premiers montrent que les composantes les plus proportionnelles sont assez éloignées des composantes principales, tout particulièrement pour les tableaux 1 et 4. Les A_k montrent des coefficients a_{k1} stables sauf une petite décroissante en $k = 4$; des a_{k2} décroissants régulièrement (de 1.383 à 0.496) et des a_{k3} croissants (de 0.860 à 1.248).

c2) L'Intrastructure Equisecteur.

Regardons l'Intrastructure Equisecteur (Fig. 6), munie des projections des axes Δv_j , associés aux composantes conjointes, et des projections des directions des compromis des variables initiales. Elle admet deux interprétations :

1) L'interprétation traditionnelle suggérée par la forme des trajectoires dans les plans principaux orientés par les directions des variables compromis. Celle-ci

Tableau 4

P 1	1.076	-0.537	-0.063	0.172	1.174	-0.253	0.008	0.027	0.892
P 2	1.057	-0.124	0.048	0.037	1.036	-0.073	0.001	0.007	0.883
P 3	1.021	0.287	0.086	-0.066	0.834	0.038	-0.000	-0.018	0.928
P 4	0.802	0.337	0.007	-0.057	0.496	0.981	0.031	-0.322	0.637
P 0	0.989	-0.009	0.020	0.021	0.885	0.173	0.010	-0.077	0.835

seuil brd 0.963 4.817

A 0 0.995 0.921 0.843

normes carrees vecteurs de C, total 17570.546 5395.623 1116.760 24082.930

jm = 1

moindres carres par tableau (Pk-Q'kAk0), total en dernier
jC=1 588.911 41.358 288.186 889.839 1808.294

q 0.891 0.426 -0.160 -0.452 0.867 -0.208 0.050 0.257 0.965

q' 1	0.720	-0.494	-0.001	0.485	0.711	0.159	-0.110	-0.115	0.987
q' 2	0.859	-0.511	0.035	0.494	0.844	0.210	-0.137	-0.164	0.977
q' 3	0.955	-0.286	0.074	0.254	0.923	0.289	-0.151	-0.257	0.955
q' 4	0.971	-0.206	0.120	-0.009	0.470	0.882	-0.238	-0.858	0.455

A 1	1.002	1.384	0.858
A 2	1.034	1.101	0.888
A 3	1.030	0.791	0.961
A 4	0.872	0.495	1.239
A 0	1.000	0.999	0.985

moindres carres par tableau (Pk-Q'kAk0), total en dernier
jC=56 32.656 14.783 28.646 8.591 84.675

$\Psi_k, k \in K$ et ϕ_s

derniere iteration pour jm,jc 1 56

moindres carres par tableau (Yk-YAk), total
jm=1 1113.847 452.034 568.151 1633.787 3767.818

$\phi_k, k \in K$ et ϕ_s

normes carrees vecteurs de C, total 17568.436 5388.616 1073.551 24030.603

jm = 2

moindres carres par tableau (Pk-Q'kAk0), total en dernier
jC=1 26.824 14.067 24.228 7.140 72.258

q 0.889 0.429 -0.161 -0.456 0.865 -0.203 0.049 0.258 0.965

q' 1	0.717	-0.497	-0.000	0.488	0.707	0.160	-0.112	-0.115	0.997
q' 2	0.856	-0.515	0.034	0.498	0.841	0.212	-0.138	-0.164	0.977
q' 3	0.954	-0.290	0.074	0.257	0.921	0.292	-0.153	-0.260	0.954
q' 4	0.970	-0.210	0.120	-0.009	0.466	0.885	-0.241	-0.860	0.450

A 1	1.001	1.384	0.861
A 2	1.034	1.101	0.885
A 3	1.031	0.792	0.960
A 4	0.872	0.496	1.247
A 0	1.000	1.000	1.000

moindres carres par tableau (Pk-Q'kAk0), total en dernier
jC=2 26.967 13.902 24.054 6.593 71.415

derniere iteration pour jm,jc 2 2

moindres carres par tableau (Yk-YAk), total
jYk=2 1134.502 457.965 565.526 1597.652 3755.645

jm = 3

moindres carres par tableau (Pk-Q'kAk0), total en dernier
jC=1 26.887 13.776 23.897 6.129 70.689

q 0.897 0.432 -0.162 -0.459 0.863 -0.209 0.049 0.260 0.964

q' 1	0.714	-0.700	-0.000	0.691	0.704	0.162	-0.113	-0.116	0.987
q' 2	0.854	-0.519	0.034	0.501	0.839	0.213	-0.139	-0.165	0.975
q' 3	0.953	-0.293	0.074	0.260	0.920	0.294	-0.154	-0.261	0.953
q' 4	0.970	-0.212	0.120	-0.009	0.464	0.886	-0.244	-0.860	0.449

A 1	1.001	1.383	0.860
A 2	1.033	1.102	0.884
A 3	1.031	0.792	0.960
A 4	0.873	0.496	1.248
A 0	1.000	1.000	1.000

moindres carres par tableau (Pk-Q'kAk0), total en dernier
jC=2 26.959 13.637 23.741 5.721 70.058

derniere iteration pour jm,jc 3 2

moindres carres par tableau (Yk-YAk), total
jYk=3 1136.582 459.383 565.930 1592.058 3754.253

A 1	1.001	1.383	0.860
A 2	1.033	1.102	0.884
A 3	1.031	0.792	0.960
A 4	0.873	0.496	1.248

normes carrees vecteurs de Y, total 14841.517 7782.892 1407.003 24031.612

erreurs relatives:et moyenne 0.0473 0.0191 0.0235 0.0662 0.0391

Q'k0 1	0.955	-0.296	0.031	0.298	0.949	-0.104	0.002	0.108	0.994
Q'k0 2	0.998	-0.070	0.003	0.370	0.996	-0.051	0.000	0.051	0.999
Q'k0 3	0.984	0.178	-0.322	-0.177	0.983	0.048	0.032	-0.044	0.999
Q'k0 4	0.964	0.267	0.003	-0.177	0.628	0.758	0.201	-0.731	0.552

3 composantes de : AFJ1 (Xk*) AFJ (Ck*) ACP (Ck) Yk=Ck0'k YAK
puis compromis AFJ1,AFJ,et Y

A5	16.52	-0.58	-0.99	15.52	1.63	1.27	14.37	6.04	1.47	14.27	-5.98	2.42	11.31	-6.03	3.28
A5	14.18	1.23	2.20	13.48	1.91	1.71	13.32	2.76	1.81	12.51	-4.89	2.81	11.68	-4.81	3.37
A5	11.29	4.33	3.53	10.81	3.23	2.43	11.16	1.39	2.61	10.59	-2.68	3.72	12.23	-3.45	3.66
A5	6.39	3.51	5.02	8.50	1.83	4.93	8.70	3.42	3.61	7.53	-3.36	5.70	9.67	-2.16	4.76
A5	12.22	2.12	2.44	12.08	2.15	2.60	11.31	-4.36	3.61						

Résultats de la dernière itération



est tout à fait comparable à l'interprétation de l'I.ACP ; nous avons choisi de ne représenter ici que 12 trajectoires.

2) Une interprétation tenant compte de la projection des axes Δv_j (dans les plans principaux) et des coefficients a_{kj} . Elle est plus délicate dans la mesure où on ne connaît pas les projections des trajectoires dans les plans $(\Delta v_j, \Delta v_{j+1})$ mais seulement les projections de ces trajectoires et de ces axes dans les plans principaux.

Sur l'axe Δv_1 , les coordonnées s'écrivent $y_i^{1k} = y_i^1 a_{k1} + \varepsilon_i^{1k}$, $k \in K$. Donc pour chaque trajectoire i , $i \in I$, aux erreurs ε_i^{1k} près, les coordonnées y_i^{1k} sont sensiblement constantes, avec une légère décroissance pour $k = 4$. D'après les projections visibles dans les plans principaux, l'axe Δv_1 oppose la catégorie *ea* à *ou* et aux catégories du tertiaire (plan 1-2) ; *ea* à *ou* (plan 1-3). Compte tenu du centrage des tableaux, la faible évolution des a_{k1} traduit, non pas la stabilité des profils de ces catégories, mais la stabilité des variances de ces profils, avec une petite décroissance pour $k = 4$.

Sur l'axe Δv_2 qui oppose les catégories *oa* à *cm*, *ou*, *cp*, la décroissance des a_{k2} , qui entraîne les trajectoires vers l'axe Δv_1 , signifie que les profils de ces catégories tendent à s'approcher du profil moyen (relatif à l'ensemble des cantons).

L'AFJ permet donc de visualiser les variables dont la variance décroît (ou s'accroît) tout en montrant les cantons qui contribuent le plus nettement à ce mouvement.

Pour l'axe Δv_3 , l'accroissement des a_{k3} pour $k = 3$ et surtout $k = 4$ montre que les catégories qui s'opposent : *oa*, *ea*, *ou*, d'une part, et toutes les autres, d'autre part, connaissent un accroissement de dispersion pour ces périodes.

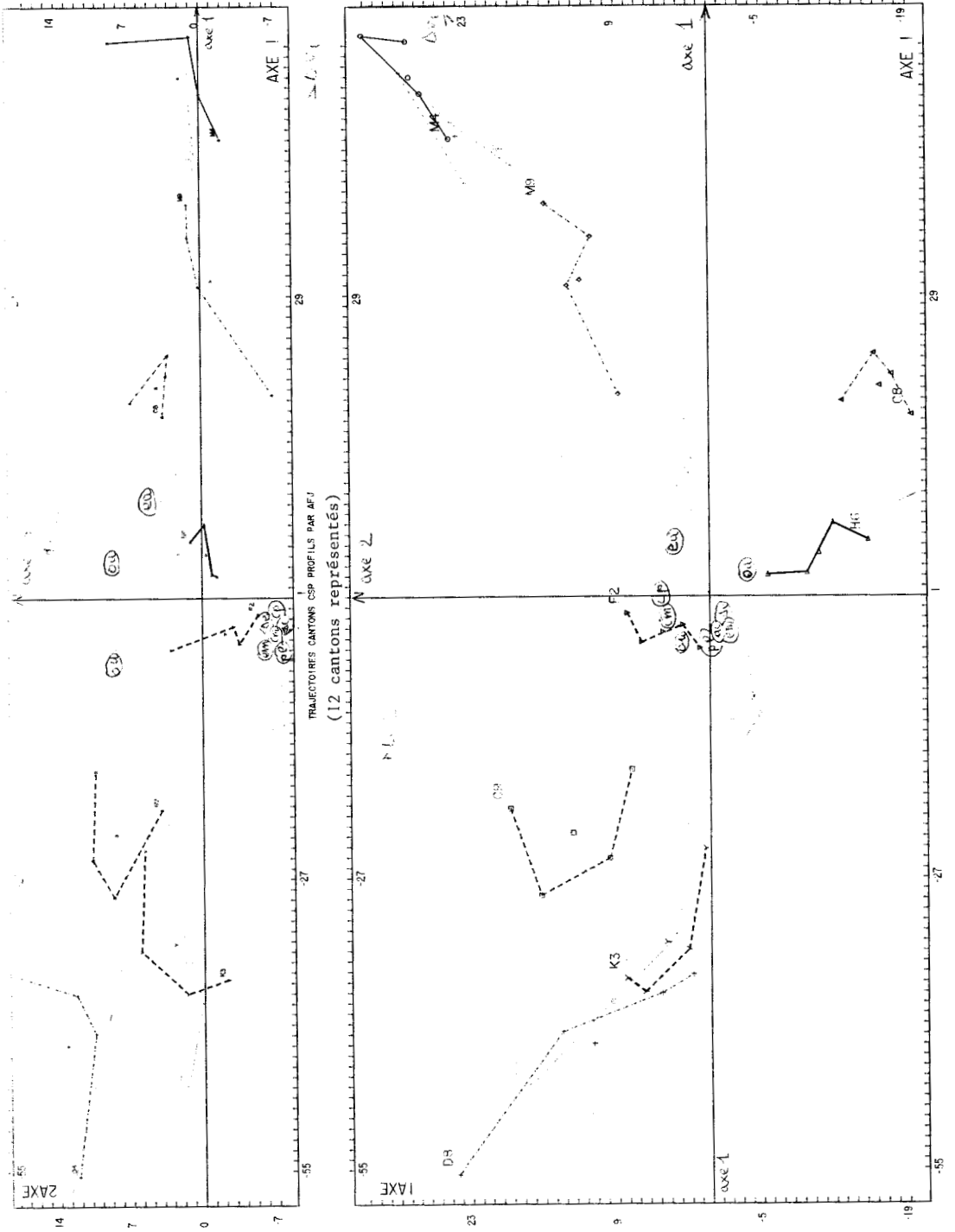
Tous ces résultats confirment (et réciproquement peuvent être affinés par) l'évolution des variances des variables :

- en hausse pour *cm*, *em*, *sc*, *pl* (par ordre décroissant)
- en baisse pour *oa*, *ou* (et *ea* très faiblement).

d) L'Intrastructure Compromis : quasi-identique à l'I.D.

L'identité de l'I.C. et l'I.D. se vérifie si $Y_k = Y A_k$, $k \in K$. Elle se vérifie approximativement si les composantes des tableaux sont quasi-proportionnelles (ici, l'erreur moyenne est 3, 9 %) et si les erreurs ε_i^{jk} telles que $y_i^{jk} = y_i^j a_{kj} + \varepsilon_i^{jk}$,

Figure 6 I.E. des cantons en profils



sont bien réparties sur tous les individus (et non concentrées sur quelques individus comme dans l'effet PLM).

En partant de l'I.E. définie par Y_k , on obtient facilement l'I.D. définie par $\hat{Y}_k = Y_k A_k$, en multipliant simplement chaque coordonnée y_i^k par a_{kj} . Cette opération peut être faite de façon approximative dans l'I.E.P. présentée à la figure précédente (fig.6). On obtient effectivement une I.D.P. très ressemblante à l'I.C. de la figure 7.

Ce cas illustre parfaitement la manière dont l'Intrastructure Compromis de l'AFJ1 introduit des déformations systématiques dont on ne peut pas connaître a priori :

- l'intensité (liée aux coefficients a_{kj})
- la direction (celle des axes Δv_j)

Ici, les déformations sont négligeables sur l'axe Δv_1 mais assez importantes sur les axes Δv_2 et Δv_3 .

e) Les cantons avec les effectifs bruts

Pour cet exemple nous laissons le soin au lecteur d'interpréter les résultats, en précisant que les figures sont données pour le plan (1-2). Pour une bonne lecture de l'I.ACP (Fig. 8), nous précisons ci-dessous les coordonnées des projections des individus-variables j.

Axes	inertie	ea	oa	ac	pl	cm	em	ou	se	cp
1	69.96	.188	.266	.281	.108	.214	.297	.822	.109	.044
2	91.74	-.574	-.731	-.103	.030	.029	.013	.353	-.002	-.005
3	95.99	.095	.097	-.114	-.382	-.509	-.590	.416	-.201	.037

L'I.E.P. de l'AFJ et l'I.C. de l'AFJ1 sont présentées sur un même graphique (Fig. 9) pour mettre en évidence le phénomène de dilatation des trajectoires de l'I.C. Elle est ici limitée à 7 cantons. Pour chacun d'eux, la trajectoire la plus courte est celle de l'I.E., la plus longue celle de l'I.C. .

On remarquera que dans cet exemple (tab. 5), l'algorithme externe de l'AFJ est parcouru 5 fois, contre 3 fois dans les exemples précédents. Cela peut tenir à différentes causes que nous préciserons au chapitre V.

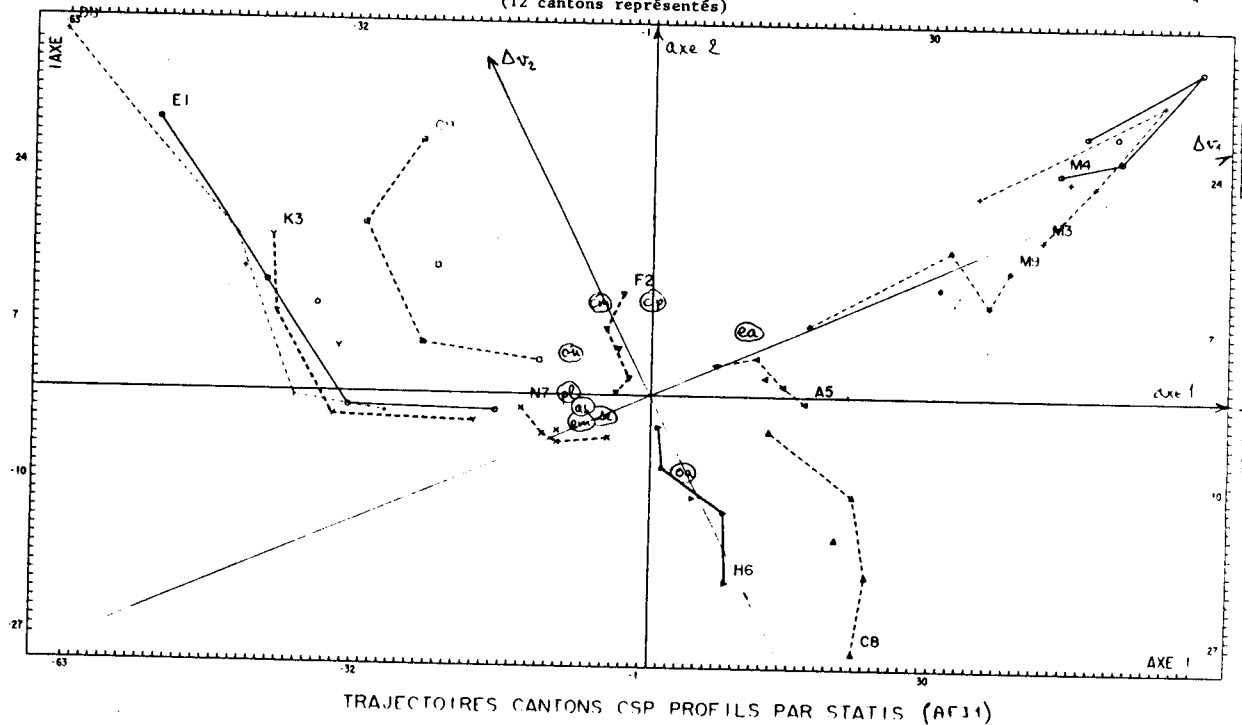
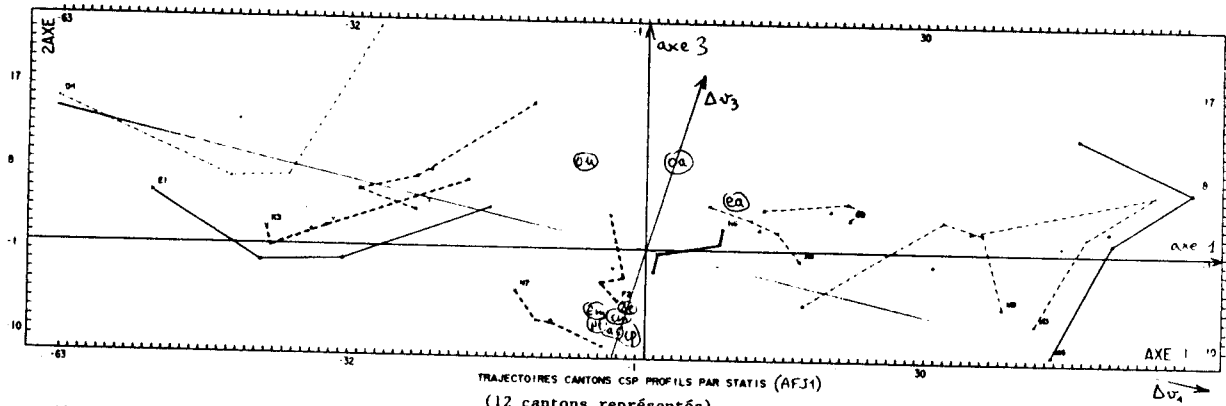


Figure 7 1.C. des cantons en profils

Tableau 5

L'AFJ sur l'exemple des cantons (effectifs bruts divisés par 100)

La boucle extérieure est parcourue 5 fois.

P 1	0.714	0.556	-0.864	-0.202	1.233	0.641	0.025	-0.025	0.282
P 2	0.788	-0.064	-0.661	0.021	1.083	-0.066	0.015	0.006	0.202
P 3	1.043	-0.187	-0.157	0.039	0.809	-0.287	0.002	0.027	0.173
P 4	1.317	0.027	1.119	0.044	0.445	-0.648	-0.036	0.318	0.516
P 0	0.965	0.083	-0.140	-0.024	0.893	-0.090	0.001	0.081	0.293

seuil b.d 0.258 1.289
A 0 0.994 0.942 0.323

normes carrees vecteurs de C, total 4696.418 1340.827 405.358 6442.602

moindres carres par tableau (Pk-Q'kAkQ), total en dernier
j_c = 1 566.929 182.145 63.783 516.826 1329.682

jm = 1

q	0.998	-0.033	-0.061	0.035	0.998	0.043	0.059	-0.045	0.997
q' 1	0.925	0.377	-0.046	-0.375	0.926	0.046	0.060	-0.025	0.998
q' 2	1.000	-0.002	-0.020	0.001	1.000	-0.021	0.020	0.021	1.000
q' 3	0.999	-0.037	0.010	0.038	0.997	-0.071	-0.007	0.071	0.997
q' 4	0.989	0.018	0.144	0.091	0.693	-0.716	-0.113	0.721	0.684

A 1	0.745	1.366	0.578
A 2	0.794	1.086	0.372
A 3	1.048	0.813	0.346
A 4	1.310	0.541	1.846
A 0	0.999	0.995	0.560

moindres carres par tableau (Pk-Q'kAkQ), total en dernier
j_c = 12 423.682 159.106 66.210 503.331 1152.329

φ_k, k ∈ K et φ_s

derniere iteration pour jm, jc 1 12

moindres carres par tableau (Yk-YAk), total
jm = 1 778.890 365.258 234.180 723.554 2101.883

φ_k, k ∈ K et φ_s

normes carrees vecteurs de C, total 4688.266 1330.317 149.188 6167.771

jm = 2

moindres carres par tableau (Pk-Q'kAkQ), total en dernier
j_c = 1 260.817 107.248 53.171 282.499 703.734

q	0.796	-0.115	-0.594	0.090	0.993	-0.072	0.598	0.004	0.801
q' 1	0.869	0.458	0.187	-0.458	0.888	-0.044	-0.186	-0.047	0.981
q' 2	0.978	0.078	0.195	-0.080	0.997	0.002	-0.194	-0.018	0.981
q' 3	0.967	0.014	0.256	-0.010	1.000	-0.018	-0.256	0.015	0.967
q' 4	0.453	0.031	0.891	0.454	0.852	-0.261	-0.767	0.523	0.372

A 1	0.911	1.366	0.289
A 2	0.970	1.080	0.284
A 3	1.250	0.811	0.434
A 4	0.817	0.556	1.910
A 0	1.007	1.002	1.018

moindres carres par tableau (Pk-Q'kAkQ), total en dernier
j_c = 33 147.309 26.652 45.575 51.628 271.164

derniere iteration pour jm, jc 2 33

moindres carres par tableau (Yk-YAk), total
jm = 2 653.449 269.712 203.807 381.891 1508.859

normes carrees vecteurs de C, total 4693.426 1336.119 307.003 6330.548

jm = 3

moindres carres par tableau (Pk-Q'kAkQ), total en dernier
j_c = 1 153.726 15.671 105.127 28.740 303.264

q	0.777	-0.130	-0.616	0.103	0.992	-0.080	0.621	-0.001	0.784
q' 1	0.869	0.477	0.128	-0.479	0.878	-0.022	-0.123	-0.042	0.997
q' 2	0.985	0.093	0.142	-0.094	0.996	-0.003	-0.142	-0.010	0.990
q' 3	0.968	0.022	0.251	-0.013	0.999	-0.037	-0.252	0.033	0.967
q' 4	0.368	0.042	0.929	0.510	0.827	-0.239	-0.778	0.561	0.283

Tableau 5 (suite)

A 1	0.957	1.366	0.206
A 2	1.005	1.078	0.208
A 3	1.264	0.811	0.407
A 4	0.690	0.561	1.936

A 0 1.001 1.000 1.001

moindres carres par tableau (Pk-Q'kAkQ),total en dernier
jc=9 136.652 10.910 116.201 17.012 280.774

derniere iteration pour jm,jc 3 9

moindres carres par tableau (Yk-YAk),total
jm=3 537.524 216.082 276.547 239.264 1269.418

normes carrees vecteurs de C,total 4691.867 1335.802 313.478 6341.147

jm = 4

moindres carres par tableau (Pk-Q'kAkQ),total en dernier
jc=1 138.245 9.804 126.792 17.744 292.585

q	0.773	-0.138	-0.619	0.109	0.990	-0.085	0.625	-0.002	0.781
q' 1	0.866	0.486	0.115	-0.487	0.873	-0.017	-0.108	-0.041	0.993
q' 2	0.987	0.102	0.128	-0.102	0.995	-0.004	-0.128	-0.009	0.992
q' 3	0.969	0.028	0.245	-0.019	0.999	-0.037	-0.246	0.031	0.967
q' 4	0.348	0.045	0.936	0.512	0.828	-0.230	-0.785	0.559	0.265

A 1	0.965	1.365	0.188
A 2	1.012	1.078	0.189
A 3	1.269	0.811	0.395
A 4	0.659	0.563	1.942

A 0 1.000 1.000 1.000

moindres carres par tableau (Pk-Q'kAkQ),total en dernier
jc=5 136.053 9.265 128.277 17.162 290.757

derniere iteration pour jm,jc 4 5

moindres carres par tableau (Yk-YAk),total
jm=4 525.795 213.382 291.916 229.829 1260.922

normes carrees vecteurs de C,total 4691.678 1336.070 314.169 6341.916

jm = 5

moindres carres par tableau (Pk-Q'kAkQ),total en dernier
jc=1 136.671 9.108 129.686 17.280 292.746

q	0.772	-0.141	-0.619	0.111	0.990	-0.087	0.625	-0.002	0.780
q' 1	0.865	0.489	0.112	-0.490	0.871	-0.016	-0.105	-0.041	0.994
q' 2	0.987	0.105	0.125	-0.105	0.994	-0.004	-0.125	-0.009	0.997
q' 3	0.970	0.030	0.243	-0.022	0.999	-0.037	-0.244	0.030	0.999
q' 4	0.344	0.046	0.938	0.510	0.829	-0.228	-0.788	0.557	0.262

A 1	0.966	1.365	0.184
A 2	1.013	1.078	0.185
A 3	1.271	0.811	0.391
A 4	0.652	0.563	1.944

A 0 1.000 1.000 1.000

moindres carres par tableau (Pk-Q'kAkQ),total en dernier
jc=2 136.421 9.048 129.819 17.264 292.552

derniere iteration pour jm,jc 5 2

moindres carres par tableau (Yk-YAk),total
jm=5 524.639 213.289 294.275 228.013 1260.217

A 1	0.966	1.365	0.184
A 2	1.013	1.078	0.185
A 3	1.271	0.811	0.391
A 4	0.652	0.563	1.944

normes carrees vecteurs de Y,total 2943.601 1370.304 2028.121 6342.027

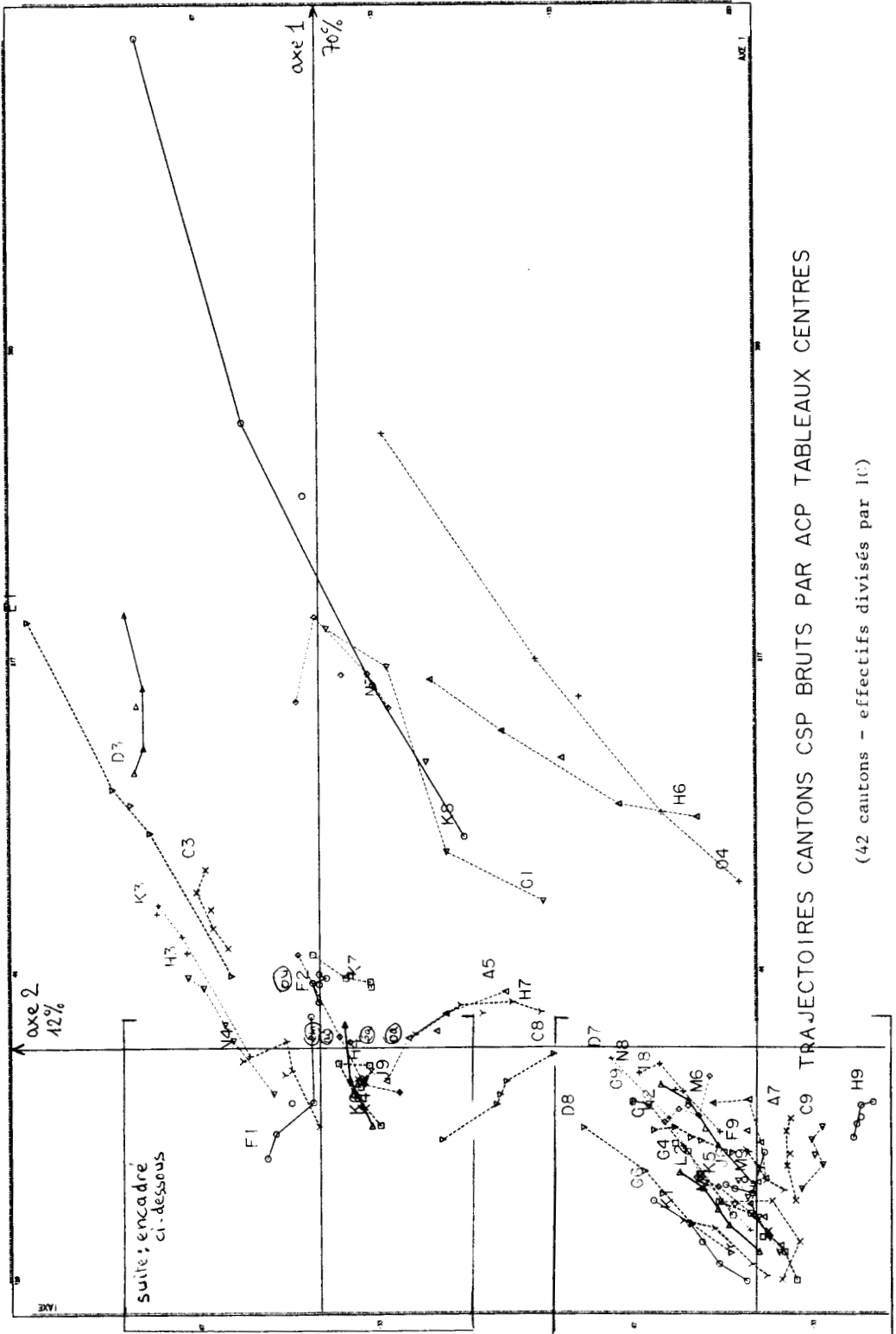
erreurs relatives:et moyenne 0.0827 0.0336 0.0464 0.0360 0.0497

C'kq 1	0.792	0.362	-0.491	-0.292	0.932	0.215	0.536	-0.027	0.844
Q'kq 2	0.852	-0.036	-0.523	0.027	0.999	-0.025	0.523	0.007	0.852
Q'kq 3	0.904	-0.107	-0.413	0.071	0.992	-0.102	0.421	0.063	0.905
Q'kq 4	0.857	-0.005	0.515	0.344	0.749	-0.566	-0.383	0.662	0.644

3 composantes de : AFJ1 (Xk[~]) ,AFJ (Ck[~]) ,ACP (Ck) ,Yk=CkQ'k ,YAk
ouits compromis AFJ1,AFJ:et Y

A5	5.58	10.69	-7.69	1.78	7.60	-10.53	9.33	4.29	-8.14	6.83	8.63	-7.12	3.13	8.45	-0.46
A5	3.06	6.32	-4.13	0.79	5.91	-6.42	3.82	6.09	-5.02	3.76	6.50	-4.53	3.29	6.67	-0.46
A5	0.85	3.83	-2.69	-0.66	4.71	-4.21	0.64	5.06	-3.79	1.44	4.96	-3.70	4.12	5.02	-0.97
A5	-3.14	1.82	-5.63	-0.78	3.27	-3.56	-2.52	4.20	0.17	1.14	3.45	-3.27	2.12	3.49	-4.84
A5	1.59	5.67	-5.04	0.28	5.37	-6.18	3.24	6.19	-2.49						
A7	-2.00	4.23	3.48	-3.45	3.31	0.09	-1.58	4.11	-1.86	-3.19	2.88	-2.09	-4.39	2.82	-0.75
A7	-3.19	3.29	2.47	-3.71	2.94	1.63	-4.12	2.80	-0.53	-4.29	2.36	-1.05	-4.61	2.23	-0.75
A7	-6.85	2.59	0.69	-6.02	2.35	2.25	-6.65	1.67	-0.35	-6.38	1.46	-2.01	-5.77	1.68	-1.59
A7	-11.26	0.60	-8.98	-7.80	1.66	-3.55	-8.52	0.57	1.80	-4.06	1.09	-7.65	-2.97	1.16	-7.89
A7	-5.93	2.68	-0.58	-5.25	2.56	0.11	-4.54	2.07	-4.06						

Figure 8 I. ACP des cantons effectifs bruts



2 fois 7 trajectoires : la plus courte est celle de l'AFJ

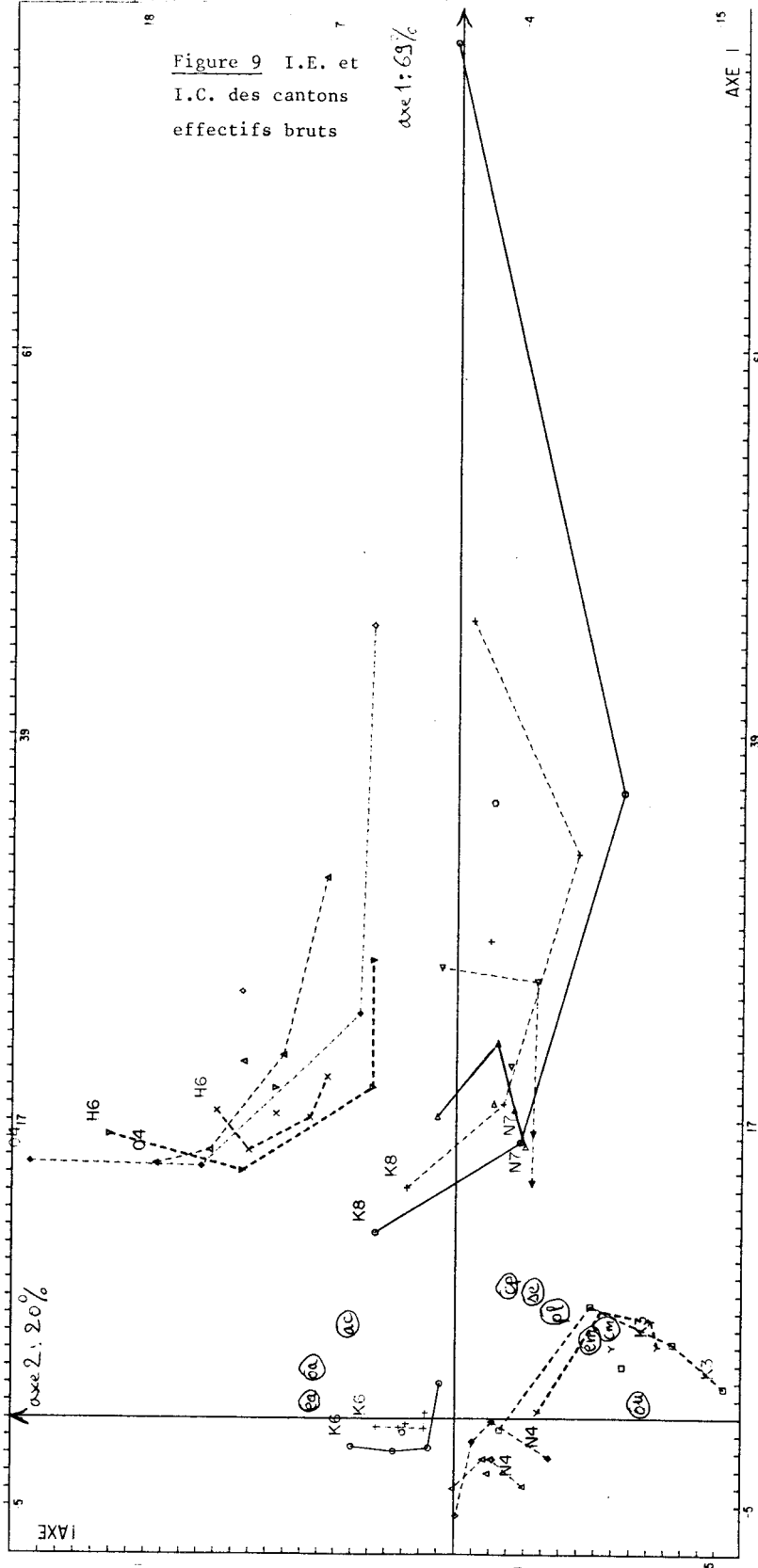


Figure 9 I.E. et I.C. des cantons effectifs bruts

TRAJECTOIRES CANTONS CSP BRUTS PAR AFJ et STATIS. (AFJt) (effectifs divisés par 100)

CHAPITRE V :

LES INTERSTRUCTURES

Nous présentons d'abord deux interstructures qui complètent celles décrites par Glaçon (1981). Ensuite nous comparons les différentes interstructures en partant d'hypothèses restrictives qui sont progressivement relâchées. Nous obtenons, en particulier, une formulation géométrique du problème INDSCAL, et une interprétation de l'interstructure de Statis qui permet de formuler un critère suffisant pour qu'une intrastructure soit significative.

5.1.- Les interstructures des A_k et des A_k^2

Supposons que les tableaux soient à composantes quasi-proportionnelles avec $Y_k = YA_k + \varepsilon_k$, $k \in K$, tels que $\sum_k \|\varepsilon_k\|^2$ soit minimale et (Y_k, I, D) soit équivalent à (X_k, M_k, D) , $k \in K$. Les tableaux Y_k (et donc les X_k) sont caractérisés par les tableaux A_k , avec une approximation ε_k . Comme $A_k = \text{diag}(a_{kj})$, $j \in [q]$, $k \in K$, on peut définir une interstructure des tableaux en positionnant dans l'espace $E_q = (R^q, I_q)$ les points s_k (situation k) de coordonnées $(a_{kj}\|y^j\|, j \in [q])$, $k \in K$.

Définition 12.- *Si les tableaux X_k admettent les composantes conjointes : $Y_k = YA_k + \varepsilon_k$, $k \in K$, une interstructure de ces tableaux s'obtient en positionnant les points s_k^1 de coordonnées $(a_{kj}\|y^j\|, j \in [q])$, $k \in K$, dans l'espace E_q . On l'appelle : interstructure des A_k .*

Sous la même hypothèse, l'interstructure des A_k^2 est obtenue par le positionnement dans E_q des points s_k^2 de coordonnées $(a_{kj}^2\|y^j\|^2, j \in [q])$, $k \in K$.

Remarques

1) Contrairement à l'interstructure de Statis qui précède son intrastructure, ces interstructures ne sont définies qu'après l'intrastructure de l'AFJ.

2) Les interstructures classiques, décrites par Glaçon (1981), seront nommées simplement : interstructures des W_k , des V_k , ou des X_k .

5.2.- Les interstructures de tableaux particuliers

a) Les tableaux proportionnels

Les triplets (X_k, M_k, D) sont tels que $C_k = a_k C$, $k \in K$.

Proposition 13.- *Pour des tableaux proportionnels :*

- 1) *les interstructures des W_k , des Λ_k et des A_k^2 coïncident.*
- 2) *il en est de même de celles des C_k et des A_k .*

Nous noterons : I, II, III, ... les axes successifs de l'interstructure des W_k , ou des Λ_k , ou des C_k ; et 1, 2, 3, ... ceux de l'interstructure des A_k , ou des A_k^2 .

■ Montrons que dans chaque cas, les objets considérés sont sur un axe et admettent la même coordonnée.

1)

- On a $W_k = a_k^2 W$, donc W et tous les W_k sont sur l'axe I, avec les coordonnées : $W_I = \|W\|_{\varphi_{DD}} = [\text{trace}(WD)^2]^{1/2} = [\sum_j \lambda_j^2]^{1/2}$, et $(W_k)_I = a_k^2 \|W\|$, $k \in K$.
- Λ_k étant la matrice des variances-covariances des colonnes des tableaux C_k , on a : $(\Lambda_k)_I = a_k^2 \|\Lambda\|$ avec $\|\Lambda\|_{\varphi_{II}} = \|W\|_{\varphi_{DD}}$ et l'interstructure des Λ_k coïncide avec la précédente.
- L'interstructure des A_k^2 est définie par les points s_k^2 de coordonnées $(a_{kj}^2 \|c^j\|^2, j \in [q])$, $k \in K$.
En considérant le vecteur $s^2 = (\|c^1\|^2, \|c^2\|^2, \dots, \|c^q\|^2)$, les points s_k^2 sont alignés sur l'axe Δs^2 , et leur coordonnée sur cet axe est : $a_k^2 \|s^2\|$, avec $\|s^2\|_{I_q} = \|\Lambda\| = \|W\| = [\sum_j \lambda_j^2]^{1/2}$.

2)

- Les tableaux C_k étant de même dimension $(n \times q)$, le produit scalaire φ_{DI} permet de définir une interstructure de ces tableaux. Ils sont alignés sur l'axe I avec : $(C_k)_I = a_k \|C\|$ et $\|C\|^2_{\varphi_{DI}} = \text{trace}(\Lambda) = \sum_j \lambda_j$.
- L'interstructure de A_k se traite comme celle des A_k^2 . En introduisant le vecteur $s^1 = (\|c^1\|, \|c^2\|, \dots, \|c^q\|)$, les points s_k^1 ont la coordonnée $a_k \|s^1\|$ sur l'axe Δs^1 , avec $\|s^1\|_{I_q} = \|C\| = [\sum_j \lambda_j]^{1/2}$. ■

Remarque :

L'interstructure des tableaux X_k ($n \times p_k$) n'est définie que si $p_k = \bar{p}$, $k \in K$ (données cubiques) de même que celle des V_k . On pourrait alors étudier s'il y a coïncidence entre les interstructures des X_k et des C_k , ou celles des V_k et des Λ_k .

b) Les tableaux à composantes principales proportionnelles

Proposition 14.- Pour des triplets (X_k, M_k, D) tels que $C_k^q = CA_k$, $k \in K$.

1) les interstructures des W_k , des Λ_k et des A_k^2 coïncident.

2) il en est de même de celles des C_k et des A_k .

■ Montrons que dans chaque cas, les objets considérés admettent le même système de coordonnées dans une base orthonormale.

1)

$$- W_k = C_k I C_k' = C A_k^2 C' = \sum_j a_{kj}^2 c^j (c^j)' = \sum_j a_{kj}^2 W_{jj},$$

avec $W_{jj} = c^j (c^j)'$, $j \in [q]$ qui vérifie les propriétés :

$$- \|W_{jj}\|^2 \varphi_{DD} = \text{trace}(c^j (c^j)' D c^j (c^j)' D) = \text{trace}((c^j)' D c^j)^2 = \lambda_j^2, j \in [q].$$

$$- \langle W_{jj}, W_{ll} \rangle \varphi_{DD} = \text{trace}(c^j (c^j)' D c^l (c^l)' D) = 0, \text{ car } (c^j)' D c^l = 0, j \in [q], l \in [q] - \{j\}$$

Dans la base orthonormale formée des vecteurs colinéaires aux W_{jj} , $j \in [q]$, du sous-espace de $(\mathbf{R}^{n^2}, D \otimes D)$ engendré par les W_k , ceux-ci ont les coordonnées : $(a_{kj}^2 \lambda_j, j \in [q])$, $k \in K$.

- $\Lambda_k = \Lambda A_k^2$ avec $\Lambda = \text{diag}(\lambda_j)$, $j \in [q]$. Donc les Λ_k admettent les coordonnées $(a_{kj}^2 \lambda_j, j \in [q])$ dans la base orthonormale formée des vecteurs que sont les matrices à éléments nuls, sauf l'élément situé en j -ième position sur la diagonale qui vaut 1, du sous-espace de $(\mathbf{R}^{q^2}, I_{q^2})$ qu'ils engendrent.

- Les coordonnées $a_{kj}^2 \|c^j\|^2 = a_{kj}^2 \lambda_j$, $j \in [q]$, des points s_k^2 , $k \in K$ sont relatifs à une base orthonormale par convention.

2) Dans la base orthonormale formée des vecteurs $c^j / \sqrt{\lambda_j}$, $j \in [q]$, du sous-espace de l'espace des variables $F = (\mathbf{R}^n, D)$ engendré par les $C_k^q = CA_k$, ceux-ci ont les coordonnées : $(a_{kj} \|c^j\|, j \in [q])$, $k \in K$. Ce sont aussi les coordonnées de l'interstructure des A_k . ■

c) Les tableaux à composantes proportionnelles

Pour des triplets tels que $Y_k = Y A_k$, $k \in K$, on peut encore écrire :
 $W_k = Y_k Y_k' = Y A_k^2 Y'$ avec $w_{jj} = y^j (y^j)'$, $j \in [q]$. On pose $\|w_{jj}\|_{\varphi_{DD}} = \|y^j\|_1^2 = \mu_j$, $j \in [q]$. Il est clair que les w_{jj} ne sont pas orthogonaux dans $(\mathbf{R}^{n^2}, D \otimes D)$ et donc les W_k sont caractérisés par des coordonnées $(a_{kj}^2 \mu_j, j \in [q])$, $k \in K$, dans une base non orthonormale. Cette interstructure ne coïncide plus avec celle des A_k^2 . On voit ainsi le caractère arbitraire de l'interstructure des A_k^2 (ou des A_k). Son intérêt est de montrer la forme très simple de l'interstructure des W_k , des Λ_k , (ou des C_k) sous les hypothèses précédentes. Remarquons qu'une simple convention sur la base permettrait encore de faire coïncider les interstructures des W_k et des A_k^2 , des C_k et des A_k . En fait, ces raisonnements valent surtout comme introduction au cas général.

5.3.- Les interstructures dans le cas général

a) Interprétation de INDSCAL dans \mathbf{R}^{n^2}

Le problème de trouver $\hat{W}_k = Y A_k^2 Y'$, $k \in K$, minimisant $\sum_k \|W_k - \hat{W}_k\|^2$ peut être posé dans $E^{n^2} = (\mathbf{R}^{n^2}, D \otimes D)$ de la façon suivante :

Ayant positionné les objets W_k dans cet espace (interstructure de Statis), on cherche le sous-espace $E_q^{n^2}$, engendré par les q "vecteurs" symétriques $w_{jj} = y^j (y^j)'$ ajustant le mieux possible, au sens des moindres carrés, ce nuage.

Les projections \hat{W}_k des W_k sur $E_q^{n^2}$ peuvent s'écrire : $\hat{W}_k = \sum_j a_{kj}^2 w_{jj}$, $k \in K$.
 En effet : $\langle W_k, w_{jj} \rangle_{\varphi_{DD}} = \text{trace}(C_k C_k' D y^j (y^j)' D) = \mu_j \sum_l (C_k^l)_j^2 = \mu_j a_{kj}^2$ car $(C_k^l)_j^2$ est l'inertie de la projection de la variable C_k^l sur l'axe Δy^j de l'espace des variables $F = (\mathbf{R}^n, D)$.

Les W_k et \hat{W}_k ont bien des composantes toutes positives dans le système de vecteurs w_{jj} , $j \in [q]$, et on a encore : $\hat{W}_k = \sum_j a_{kj}^2 w_{jj} = Y A_k^2 Y'$, $k \in K$.

Donc ce problème géométrique coïncide avec le problème INDSCAL qui minimise $\sum_k \|Y_k Y_k' - Y A_k^2 Y'\|^2_{\varphi_{DD}}$. Il est proche de l'AFJ qui minimise $\sum_k \|Y_k - Y A_k\|^2_{\varphi_{DI}}$. Les solutions Y et A_k , $k \in K$, sont voisines pour les deux cas même si elles ne sont donc pas identiques. Cette analyse permet d'associer à INDSCAL une interstructure qu'on nommera **interstructure des \hat{W}_k** ayant les particularités suivantes :

- elle est l'interstructure contenue dans un sous-espace de type $E_q^{n^2}$, la plus proche, au sens des moindres carrés, de celle des W_k .
- elle est comparable à celle des A_k^2 avec deux différences :
 - les a_{kj} et les y^j sont voisins mais différents pour les deux méthodes,
 - les axes sont supposés orthogonaux pour les A_k^2 , pas pour les \hat{W}_k .

b) L'interstructure de l'AFM

Dans l'interstructure des \hat{W}_k , on a cherché q "vecteurs" symétriques w_{jj} tels que le sous-espace $E_q^{n^2}$ soit le plus proche des W_k , ou ce qui revient au même, tels que la somme des carrés des normes des projetés soit maximale.

On peut modifier ce critère pour maximiser la somme des normes des projetés soit : $\sum_k \|\hat{W}_k\|^2 = \sum_k \sum_j \langle W_k, w_{jj} \rangle \varphi_{DD}$, avec $\langle W_k, w_{jj} \rangle = \mu_j \sum_l (C_k^l)_j^2$.

On sait que les variables y^j qui maximisent, dans $F = (\mathbf{R}^n, D)$, la somme des projections sur Δy^j de l'ensemble des variables du tableau X_ρ , sont successivement les composantes principales C_0 du tableau X_ρ .

Avec ce critère, Escofier et Pagès (1988) proposent de considérer l'interstructure des \hat{W}_k obtenus en projetant les W_k sur le sous-espace $E_c^{n^2}$ de E^{n^2} engendré par les "vecteurs symétriques" $W_{jj} = c^j(c^j)'$, $j \in [q]$.

L'interstructure des \hat{W}_k (de l'AFM) s'écrit $\hat{W}_k = \sum_j a_{kj}^2 W_{jj}$, $k \in K$, où chaque a_{kj}^2/m s'interprète comme la contribution du paquet J_k de variables à l'inertie de la composante c^j dans F . Les auteurs soulignent la concordance de cette interstructure rapportée à la base des $W_{jj} = c^j(c^j)'$, avec l'intrastructure de l'AFJ1 obtenue en projetant les nuages sur la base principale $(u_j)_{j \in [r]}$ associée à ces composantes principales. Comme on l'a vu cette concordance ne conduit en rien à une intrastructure exacte et significative. De même les coefficients a_{kj}^2 de cette interstructure (tels que $a_{kj}^2 = \|P_k^j\|^2$ pour les tableaux P_k introduits en 3.1.a. n'ont pas de relation précise avec les a_{kj} de l'AFJ ou de INDSCAL. Cette interstructure a donc peu d'intérêt alors que l'interstructure des W_k de la méthode Statis peut fournir un critère simple d'existence d'une intrastructure significative.

c) L'interstructure de Statis

Les raisonnements qui vont suivre contiennent un certain nombre d'approximations. Aussi les résultats présentés, concernant l'interstructure de Statis et

son interprétation, auront un caractère indicatif jusqu'à ce qu'une recherche plus approfondie vienne les valider. L'interstructure des W_k s'obtient par une classique ACP du nuage des points W_k , considérés comme des variables de \mathbf{R}^{n^2} . Comme le nombre m de tableaux est en général assez petit (ex : $3 < m < 10$), ce nuage engendre un sous-espace de petite dimension et les premiers axes portent des parts importantes d'inertie. On a une très bonne représentation des W_k avec les 2 ou 3 premiers plans d'inertie. En particulier, on a : $(W_k)_I > 0$, $k \in K$, et l'axe I porte toujours beaucoup d'inertie. (Cette particularité disparaît dans l'interstructure centrée). Adoptons les hypothèses suivantes :

1) Les tableaux sont à composantes quasi-proportionnelles donc l'intrastructure est significative et on a : $Y_k = Y A_k + \varepsilon_k$ avec $\sum_k \|\varepsilon_k\|^2$ minimale.

2) La quantité $\sum_k \|W_k - Y A_k^2 Y'\|^2$ calculée pour l'AFJ est proche du minimum atteint par INDSCAL pour $\sum_k \|W_k - \hat{W}_k\|^2$ et donc les W_k sont proches de leur estimation par $Y A_k^2 Y'$ calculée par l'AFJ.

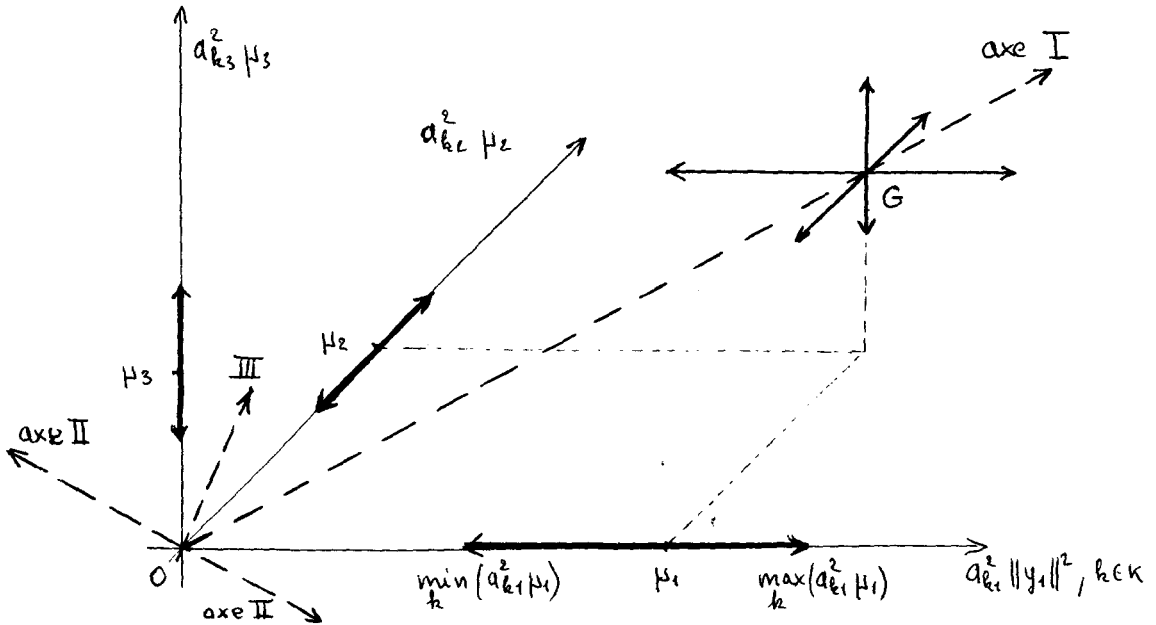
3) Les vecteurs $w_{jj} = y^j (y^j)'$ sont proches de l'orthogonalité dans $E^{n^2} = (\mathbf{R}^{n^2}, D \otimes D)$. De plus, les $\|y^j\|^2 = \mu_j$ étant rangés par ordre décroissant, ils deviennent vite assez petits pour être négligés. Ainsi une représentation graphique dans \mathbf{R}^3 constitue une approximation raisonnable (elle permet d'imaginer ce qui se passe dans des espaces de plus grande dimension s'ils sont nécessaires).

c1) Quelques remarques

1) La figure 10 qui présente l'interstructure des A_k^2 (de coordonnées $(a_{kj}^2 \mu_j, j \in [3])$, dans \mathbf{R}^3), $k \in K$, constitue une bonne approximation de l'interstructure des W_k . Nous admettrons donc que les résultats pour la première sont valables pour la seconde.

2) Le nuage des W_k est contenu dans le parallélépipède symbolisé par les 3 flèches concourantes en G . Elles ont leurs extrémités au centre des faces et leur longueur est $[\text{Max}_k(a_{kj}^2 \mu_j) - \text{Min}_k(a_{kj}^2 \mu_j)]$, $j \in [3]$.

3) La normalisation $\sum_k a_{kj}^2 = m$, $j \in [q]$ permet d'écrire : $a_{kj}^2 \in [0, m]$, $j \in [q]$, $k \in K$. Si de plus les a_{kj}^2 sont proches de 1 (par ex : $a_{kj}^2 \in [0.5, 1.5]$) on peut considérer que l'axe I de l'interstructure passe par le point G qui est proche du centre de gravité des W_k .

Fig. 10 : Les axes I, II, III de l'interstructure des W_k 

4) La dispersion des W_k relative aux axes I, II, III, ... de l'interstructure traduit donc :

- la dispersion induite par les a_{kj} qui exprime la (quasi-) proportionalité des composantes,
- une dispersion résiduelle qui devient éventuellement prépondérante lorsque les a_{kj} sont voisins de 1.

5) Nous avons préconisé (cf. 3.3.b.) la normalisation

$$\|X_k\|^2 \varphi_{DM} = \|W_k\| \varphi_{DD} = \text{cste}, \text{ pour les raisons suivantes :}$$

- Si les tableaux sont quasi-proportionnels ($C_k = a_k C + \varepsilon_k$), elle revient à les traiter comme des tableaux quasi-égaux. Alors toutes les infrastructures sont presque exactes (Annexe IV) et l'algorithme de l'AFJ converge extrêmement rapidement.

- Si les tableaux sont à composantes quasi-proportionnelles ($Y_k = Y A_k + \varepsilon_k$), l'avantage de cette normalisation n'est pas prouvé, sauf s'il existe une certaine proportionalité des tableaux eux-mêmes. La normalisation systématique a donc

l'avantage de traiter cette éventualité et d'améliorer la vitesse de convergence de l'algorithme.

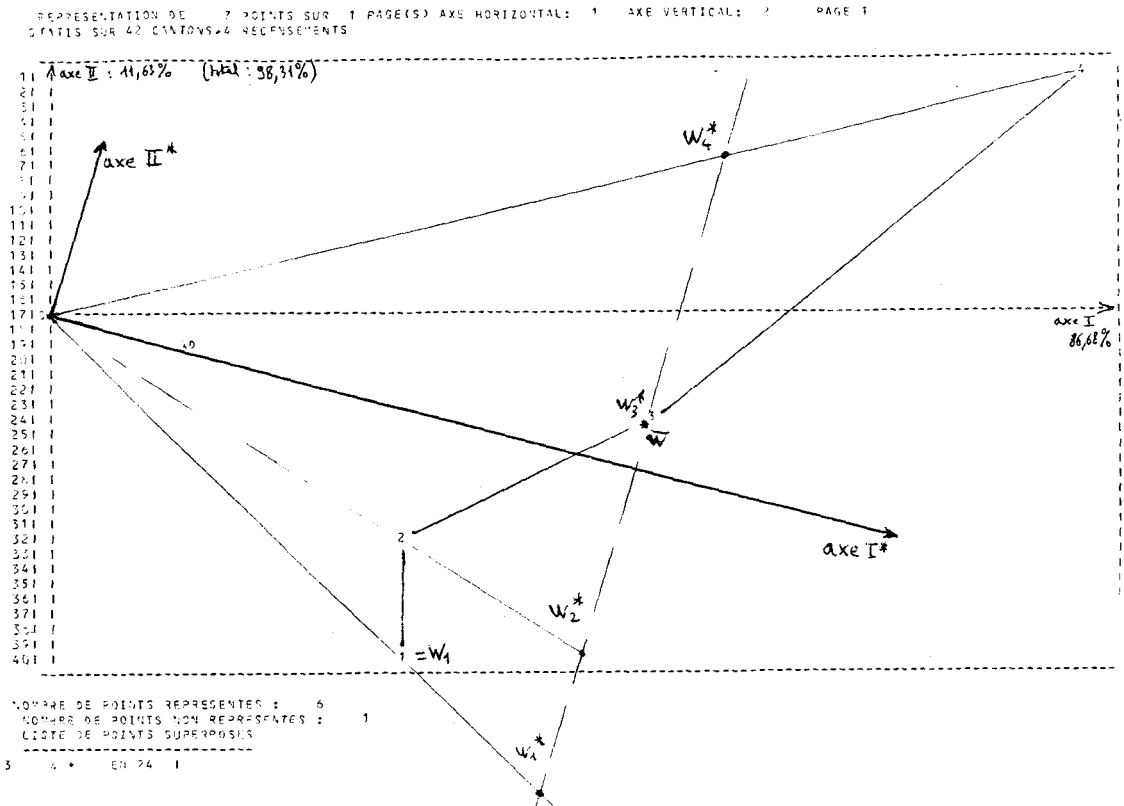
- Enfin, il est évident que dans la minimisation du critère $\Phi = \sum_k \|Y_k - YA_k\|^2$, les tableaux Y_k de norme élevée ont un "poids" plus grand que les autres. Cela donne une deuxième raison de normaliser les données.

Cette normalisation se traduit graphiquement par le fait que les W_k sont sur l'hypersphère de centre 0 et de rayon $r = (\sum_j \mu_j)^{1/2}$.

c2) L'interprétation des axes de l'interstructure

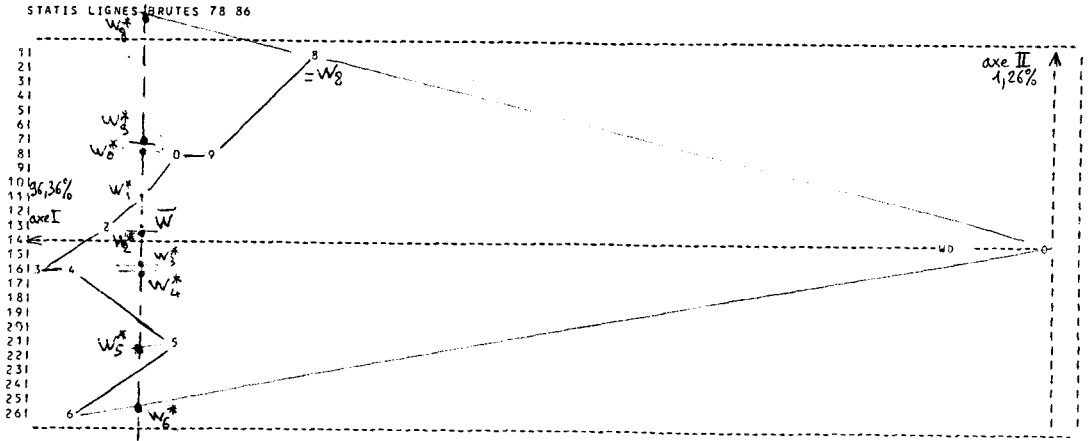
Une autre normalisation, très proche de la précédente lorsque les a_{kj} sont voisins de 1 (ce qui donne un nuage des W_k suffisamment petit), consiste à prendre W_k^* proportionnel à W_k tel que : $(W_k^*)_I = \bar{W}_I$, $k \in K$, avec : $\bar{W} = \frac{1}{m} \sum_k W_k$. Cela peut provoquer un déplacement de l'axe I de l'interstructure (Fig. 11), et on peut réitérer ce calcul jusqu' à ce que tous les W_k^* aient bien la même première composante sur l'axe I, noté I*, de leur interstructure.

Fig. 11 : Cantons effectifs bruts. Interstructures des W_k et des W_k^*



On peut aussi négliger ces itérations car elles changent très peu les résultats surtout pour des données déjà normalisées au départ. Alors on présentera (Fig. 12 ou 13) l'interstructure normalisée (des W_k^*) dans les axes de l'interstructure des W_k .

Fig. 12 : Accidents 1978 à 1986. Interstructures des W_k et des W_k^*



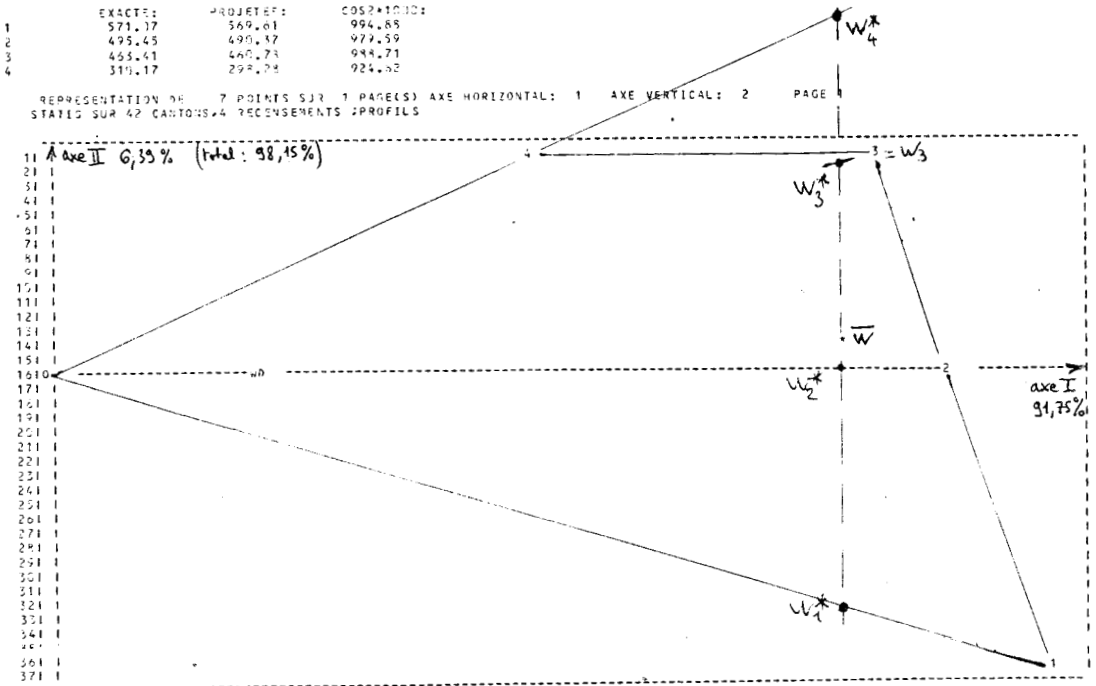
Les axes de l'interstructure s'interprètent ainsi :

- **Axe I : axe de normalisation**

L'interstructure ayant été calculée (sur les tableaux normalisés lorsque les ensembles J_k ne contiennent pas les mêmes variables), on (re-)normalise les données en prémultipliant les tableaux par $\alpha_k = (\bar{W}_I / (W_k)_I)^{1/2}$. Pour les données cubiques, on peut prendre des poids ρ_k proportionnels à α_k pour calculer le compromis de l'AFJ1 : $W_0 = \sum_k \rho_k^2 W_k$. Alors les W_k^* vérifient $(W_k^*)_I = \bar{W}_I$.

- **Axe II : axe d'évolution**

Sous la condition précédente, l'axe II* devient la direction de plus grande variabilité du nuage des W_k^* . Sur les exemples traités, l'ordre induit par les coordonnées des W_k^* sur cet axe coïncide exactement avec l'ordre chronologique et il nous paraît raisonnable de le considérer comme l'axe d'évolution. Cet ordre ne coïncide pas nécessairement avec l'ordre pour lequel la somme des distances entre W_k^* successifs est minimale. Cette notion plus complexe nous paraît moins pertinente que la précédente pour définir une évolution des situations lorsque celle-ci n'est pas donnée.

Fig. 13 : Cantons en profils. interstructures des W_k et des W_k^* 

c3) Faisabilité de l'intrastructure

La dispersion des W_k dans le sens de l'axe I ne gêne absolument pas la comparabilité des tableaux car, après normalisation, cette dispersion disparaît du nuage des W_k^* .

Si les a_{kj} sont voisins de l'unité ($a_{kj}^2 \in [0.5, 1.5]$, au moins pour les premières valeurs de j), la figure 10 "montre" que l'axe I porte une grande part de l'inertie du nuage des W_k . Dans ce cas qu'on rencontre généralement, cette part d'inertie ne diminue que si on s'écarte de l'hypothèse de composantes quasi-proportionnelles.

Conséquence :

Un critère suffisant pour que les tableaux soient à composantes quasi-proportionnelles est que l'axe I de l'interstructure porte une part p_I de l'inertie du nuage des W_k (ou W_k^) vérifiant (par exemple) : $p_I > 70\%$. L'inégalité $p_I < 70\%$ indique que les premiers a_{kj} sont éloignés de l'unité ou, que les composantes ne sont pas quasi-proportionnelles.*

A titre indicatif nous rappelons dans le tableau ci-dessous les valeurs de p_I (inertie portée par l'axe I de l'interstructure), de p_3 (part d'inertie p_s du nuage compromis portée par les trois premiers axes principaux) et l'erreur moyenne e pour nos trois exemples.

(en %)	Accidents	Cantons profils	Cantons bruts
p_I	96,36	91,75	86,68
p_3	97,73	94,58	94,80
e	2,30	3,90	4,97

Dans les trois cas, le critère suffisant relatif à p_I montre clairement la faisabilité de l'intrastructure. On voit que sa valeur plus basse pour les cantons bruts tient à une erreur moyenne plus élevée et surtout au fait que certains a_{kj} sont élevés ($a_{k3} \in [0.18, 1.94]$). Ils sont en fait des a_{k2} car $\|y^3\| > \|y^2\|$.

Le critère nécessaire relatif à p_3 indique aussi la faisabilité de cette intrastructure, y compris en se limitant à $s = 3$ composantes (on a $p_s > 95\%$ ou presque, de même pour les parts p_s^k des différents nuages).

Pour ces trois exemples, les résultats de l'AFJ sont excellents puisque nous avons une intrastructure exacte (à 95 % pour $s = 3$ composantes) et parfaitement significative puisque l'erreur moyenne e est inférieure à 5 %.

CONCLUSION

a) La pratique de l'AFJ

Pour terminer cette étude, nous résumons les étapes de l'Analyse Factorielle conJointe, conçue comme méthode de détermination d'une inrastructure exacte et significative si elle existe.

1) La normalisation des tableaux

Les tableaux sont centrés. Pour les données cubiques, on réduit les variables puis on multiplie les tableaux par \mathcal{I}/p_k pour avoir : $\|X_k\| = \mathcal{I}$, $k \in K$. Pour les données non-cubiques, on ne réduit pas les variables lorsqu'on veut rendre compte de l'effet de l'évolution de leur dispersion. Alors on adopte les poids $\rho_k = \alpha/\|X_k\|$, ou $\rho_k = \beta/(W_k)_I$ définis par l'interstructure. (α et β sont des constantes de normalisation).

2) Faisabilité de l'inrastructure

21) Le critère suffisant (l'interstructure)

- Soit p_I la part d'inertie portée par l'axe I de l'interstructure. L'inrastructure est réalisable sous le critère suffisant que $p_I > 70\%$. Elle se traduit au plan graphique par un cône dans le plan (I, II), dont l'angle au sommet est inférieur à environ 80° .
- L'interstructure donne aussi :
 - Les poids $\rho_k = \frac{\beta}{(W_k)_I}$, $k \in K$, pour les données cubiques
 - Un ordre d'évolution des situations visible sur l'axe II* de l'interstructure des W_k^* .

22) Le critère nécessaire (le compromis)

La décomposition en valeur singulière du compromis :

$X_\rho = (\rho_1 X_1 \quad \dots \quad \rho_m X_m)$ s'écrit : $X = \Gamma \Delta U'$. Avec $\Delta = \text{diag}(\sqrt{\lambda_i})$, $q = \sup(q_k)$, $r = \text{rang}(X_\rho)$ et $p = (\sum_{j=1}^q \lambda_j) / (\sum_{j=1}^r \lambda_j)$, le critère $p > p_0$ (par exemple $p_0 = 80\%$) est nécessaire pour avoir des composantes quasi-proportionnelles.

L'inrastructure est d'autant plus significative que ces deux critères sont éloignés de ces bornes. Dans le cas contraire, l'Analyse Canonique Généralisée et l'AFD associée permettent de déterminer les variables et les individus qui

détériorent le plus l'intrastructure. Ils peuvent ainsi être mis en supplémentaires dans une intrastructure qui retrouve sa signification.

3) L'AFJ détermine successivement :

- Les s.v.d. : $X_k = \Gamma_k \Delta_k U_k'$ donnant $C_k = \Gamma_k \Delta_k$ et $\Delta_k = \text{diag}(\sqrt{\lambda_i^k})$, $k \in K$.
- Le nombre minimal s de composantes à conserver. Si pour s composantes du compromis, on a $p_s > p_o^s$ et de même pour chaque tableau $p_s^k > p_o^s$ ($p_o^s = 95\%$ semble un bon choix), on se limitera à s composantes dans les tableaux C et C_k , $k \in K$.
- Les tableaux $P_k = C_k' D C \Lambda^{-1}$, $k \in K$.
- Les tableaux Y , Q , Q_k , A_k , ($k \in K$) calculés par l'algorithme qui détermine aussi les normes carrées : $\|P_k - Q_k' A_k Q\|^2$, $\|Y_k - Y A_k\|^2$, $k \in K$, et l'erreur moyenne $e = \sum_k \|\varepsilon_k\|^2 / m \|Y\|^2$.

4) L'interprétation des résultats.

Elle vient d'être présentée sur des exemples et nous n'en reprenons que deux aspects.

- Pour l'intrastructure complète, l'inertie représentée de chaque nuage vaut $p_s^k (> 95\%)$ et de même, celle du nuage compromis vaut p_s . Avec $s = q$ axes cette intrastructure est parfaitement exacte. En revanche on ne connaît pas la part d'inertie d'un nuage portée par un axe. Si les composantes principales sont proportionnelles, la part de l'inertie du nuage compromis portée par chacun de ses axes égale la part d'inertie de la réunion des nuages portée par cet axe.
- L'ACP du tableau compromis X_p fournit une représentation des variables de l'ensemble $J = \bigcup J_k$. Les composantes principales du compromis sont sur les axes de ce positionnement. On peut y ajouter les composantes principales c_k^j des différents tableaux et les variables y^j les plus proches de la proportionalité. Ainsi les caractéristiques des tableaux apparaissent dans l'intrastructure (espace des individus), et de manière duale dans l'espace des variables avec cette intrastructure des variables.

b) L'AFJ et l'analyse des données

Ce travail apporte les trois résultats suivants :

1) Avec l'AFJ on dispose d'une méthode donnant une intrastructure exacte et significative. "Exacte" indique la possibilité d'obtenir des images euclidiennes équivalentes. "Significative" précise que le positionnement simultané respecte "le mieux possible au sens des moindres carrés" les propriétés de dualité. Par ailleurs, nous avons indiqué que les deux premières étapes de la méthode Statis-l'interstructure et le compromis- donnent respectivement un critère suffisant et un critère nécessaire pour que l'intrastructure soit significative.

2) Exceptons INDESCAL qui est une méthode des moindres carrés dans \mathbb{R}^{n^2} , et les méthodes de type ACP-3 modes susceptibles de réduire le mode des situations ; les autres méthodes de l'analyse conjointe de tableaux de données ont été présentées comme techniques de \sqrt{m} -projection sur un certain sous-espace, des nuages d'individus $\tilde{N}_k, k \in K$, positionnés dans l'espace somme directe orthogonale des espaces initiaux. Un lien très simple entre les méthodes d'origine française ou anglo-saxonne apparaît ainsi très clairement.

3) Cette présentation a permis de dégager des propriétés duales sur les espaces des individus et sur l'espace des variables :

- Pour que l'intrastructure exacte soit significative il faut que les sous-espaces F_k engendrés respectivement par les variables des tableaux X_k soient inclus dans l'un d'eux. Elle l'est d'autant moins qu'on s'écarte de cette situation.
- L'Intrastructure Equisecteur correspond aux composantes (conjointes) les plus proches de la proportionalité au sens des moindres carrés.
- L'Intrastructure Procuste correspond aux composantes (procustes) les plus proches de l'égalité à des composantes principales, au sens des moindres carrés.
- L'Intrastructure Compromis (non exacte) correspond aux variables canoniques au sens de la liaison $L^2(y, X_k)$. (Escofier et Pagès 1988).
- L'Intrastructure Discriminante (non exacte) correspond aux variables (et composantes) canoniques de l'ACG au sens de Carroll (Annexe II).

Nous avons limité cet exposé au cas des variables quantitatives. Rien n'interdit de traiter par l'AFJ des familles de tableaux de variables qualitatives remplacées par les indicatrices des modalités, ou des tableaux de contingence. Mais il nous paraît nécessaire de formuler précisément la signification d'une telle méthode pour interpréter correctement ses résultats. Ce sera l'objet d'une prochaine recherche.

ANNEXE I : Dualité, Biplot en ACP, application à l'AFC

L'AFJ a été définie par application des propriétés de dualité à l'analyse conjointe. Ici, nous reprenons les propriétés relatives à un seul tableau. La dualité se traduit par une certaine symétrie des traitements effectués sur les individus ou les variables par l'ACP. La s.v.d. du tableau X s'écrivant $X = \Gamma\Delta U'$, ces traitements sont résumés par les deux schémas suivants :

—pour les individus

$$\begin{array}{ccc}
 E = \mathbf{R}_{(e_j)}^p & \xleftarrow{X'} & F^* & & E_q = \mathbf{R}_{(u_j)}^q & \xleftarrow{C'} & F^* \\
 M \downarrow \uparrow V & & W \downarrow \uparrow D & \iff & I \downarrow \uparrow \Lambda = \Delta^2 & & W \downarrow \uparrow D \\
 E^* & \xrightarrow{X} & F = \mathbf{R}_{(f_i)}^n & & E_q^* & \xrightarrow{C = \Gamma\Delta} & F
 \end{array}$$

—pour les variables

$$\begin{array}{ccc}
 E = \mathbf{R}_{(e_j)}^p & \xleftarrow{X'} & F^* & & E & \xleftarrow{Z'} & F_q^* \\
 M \downarrow \uparrow V & & W \downarrow \uparrow D & \iff & M \downarrow \uparrow V & & \Delta^2 \downarrow \uparrow I \\
 E^* & \xrightarrow{X} & F = \mathbf{R}_{(f_i)}^n & & E^* & \xrightarrow{Z = \Delta U'} & F_q = \mathbf{R}_{(\gamma^i)}^q
 \end{array}$$

Par ailleurs, Cailliez, Pagès (1976) montrent que la dualité fonde les propriétés suivantes :

- L'individu $i \in I$ est à la fois vecteur x_i de E , vecteur f^{i*} de F^* , axe Δf^i de F .
- La variable $j \in J$ est à la fois vecteur x^j de F , vecteur e_j^* de E^* , axe Δe_j de E .

a) Le biplot en ACP

L'espace E (resp. F) contient un positionnement des individus (resp. des variables), mais aussi une indication des directions Δe_j des variables (resp. Δf^i pour celles des individus) qui peuvent être projetées comme les individus (resp. les variables) sur les divers plans principaux. Pour déterminer ces projections, il est intéressant de considérer les vecteurs e_j de la base canonique de E , qui définissent les axes Δe_j (nous laissons maintenant le raisonnement dual relatif à F).

Chaque e_j peut être considéré comme un individu de l'espace E . C'est l'individu pour lequel toutes les variables observées sont nulles (après centrage), sauf la variable j qui vaut 1. Il est donc légitime d'appeler e_j : **l'individu-variable j** . Par projection, il permet d'obtenir une représentation simultanée des individus et

des directions des variables initiales. La lecture et l'orientation des plans principaux en seront largement facilitées.

Les composantes principales qui forment les colonnes de C sont les composantes des individus dans la base $(u_j)_{j \in [q]}$ de E_q , notée (u_j) , dont les vecteurs sont les colonnes de $U(p \times q)$. En effet, les colonnes de $C^c = (C \vdots 0)$ sont les composantes des individus dans la base M -orthonormale $(u_j)_{j \in [p]}$ de E , notée (u_j^c) , dont les vecteurs sont les colonnes de $\tilde{U}(p \times p)$. Les composantes de la base canonique $(e_j)_{j \in [p]}$ de E , dans la base (u_j^c) sont les colonnes de \tilde{U}^{-1} . Et $\tilde{U}^{-1}M\tilde{U} = I \Rightarrow \tilde{U}^{-1} = \tilde{U}'M$. Donc les q composantes des vecteurs de la base canonique dans la base (u_j) sont les colonnes de $\Upsilon = U'M(q \times p)$ et les directions des p variables représentées dans l'espace E_q passent par les points 0_E et $\Upsilon^j, j \in J$. De même dans les plans principaux avec les projections $\hat{\Upsilon}^j$ des Υ^j .

Pour les logiciels qui ne donnent pas ces directions, il est facile de les retrouver à partir du positionnement de la variable j , défini par la colonne j du tableau $Z = \Delta U'$, soit : $Z^j = \Delta U'^j$. Sous l'hypothèse $M = \text{diag}(m_j)$, on a : $\Upsilon_i^j = \frac{1}{\sqrt{\lambda_i}} m_j Z_i^j, i \in [q], j \in J$. En particulier, pour une métrique identité et pour le premier plan principal d'inertie : $\Upsilon_1^j = \frac{1}{\sqrt{\lambda_1}} Z_1^j, \Upsilon_2^j = \frac{1}{\sqrt{\lambda_2}} Z_2^j, j \in J$. On a aussi une indication sur la proximité de l'individu-variable j avec le plan principal. Celle-ci peut être mesurée par $\cos \theta_j = \|\hat{\Upsilon}^j\|$ directement lisible sur ce plan. Les variables normées $Z^j/\|Z^j\|$ qui figurent dans le cercle des corrélations ne donnent pas immédiatement ce dernier résultat. Les logiciels d'ACP devraient donc faire figurer dans les plans principaux du nuage des individus, les points $\hat{\Upsilon}^j, j \in J$ qui déterminent :

- Les projections des directions Δe_j des variables initiales,
- Le cosinus de l'angle θ_j entre ces directions et ces plans : $\cos \theta_j = \|\hat{\Upsilon}^j\|$.

Cette conception du biplot fondée sur les propriétés de dualité diffère de la conception développée dans la littérature de langue anglaise (Jolliffe 1985). Dans celle-ci, le biplot se justifie par des raisons pratiques (les renseignements divers qu'on en tire) et non théoriques. Comme le note Saporta (1990), certaines versions ne reçoivent pas de justification.

b) Application à l'AFC

Soit $N(m \times p)$ un tableau de contingence répartissant n individus selon les deux caractères 1 et 2 ayant respectivement m et p modalités. L'AFC est souvent

présentée comme superposition des résultats des ACP (centrées ou non car il y a équivalence) des profils des modalités de chaque caractère. On fait donc les ACP des triplets $(X = D_m^{-1}F, D_p^{-1}, D_m)$, et $(Y = D_p^{-1}F', D_m^{-1}, D_p)$, avec :

$F = \frac{1}{n}N$, tableau des fréquences (f_{ij}) , $i \in [m]$, $j \in [p]$,

$D_m = \text{diag}(f_{i.})$, $D_p = \text{diag}(f_{.j})$, matrices diagonales des fréquences marginales.

Nous allons détailler l'ACP du premier triplet en introduisant dans l'espace des "individus" (les m modalités du caractère 1), la représentation des individus-variables j . Le schéma de dualité s'écrit :

$$\begin{array}{ccc} E = \mathbf{R}_{(e_j)}^p & \xleftarrow{X'} & F^* \\ D_p^{-1} \left\{ \begin{array}{c} \uparrow \\ V \\ \downarrow \end{array} \right. & & W \left\{ \begin{array}{c} \uparrow \\ D_m \\ \downarrow \end{array} \right. \\ E^* & \xrightarrow{X} & F = \mathbf{R}_{(f_i)}^m \end{array}$$

Les individus sont représentés par les coordonnées contenues dans les lignes de $C = (c^1 \quad \vdots \quad c^2 \quad \vdots \quad \dots \quad \vdots \quad c^q)$ où chaque c^α est vecteur propre de WD_m , D_m -normé à $\sqrt{\lambda_\alpha}$.

Et on a : $WD_m = X'D_p^{-1}XD_m = D_m^{-1}FD_p^{-1}F'D_m^{-1}D_m = XY$.

Les individus-variables j ont leurs coordonnées définies par les colonnes de $\Upsilon = U'M$ avec $U = (u_1 \quad \vdots \quad u_2 \quad \vdots \quad \dots \quad \vdots \quad u_q)$, où chaque u_α est vecteur propre de VD_p^{-1} , D_p^{-1} -normé à 1.

Et on a : $VD_p^{-1} = X'D_m^{-1}XD_p^{-1} = F'D_m^{-1}D_mD_m^{-1}FD_p^{-1}D_m = X'Y'$.

Ces coordonnées sont aussi les lignes de la matrice des facteurs $B = MU$, avec $B = (b^1 \quad \vdots \quad b^2 \quad \vdots \quad \dots \quad \vdots \quad b^q)$, où chaque b^α est vecteur propre de $D_p^{-1}V = YX$, D_p -normé à 1 (ou V -normé à $\sqrt{\lambda_\alpha}$).

On en déduit les représentations simultanées des modalités i du caractère 1 et j du caractère 2. Dans le système des axes principaux d'inertie du nuage des profils $f_{j/i}$, $i \in [m]$:

- Les lignes de C ($m \times q$) sont les coordonnées des modalités i ,
- Les lignes de B ($p \times q$) sont les coordonnées des modalités j .

L'égalité $C = XB$ s'écrit, pour chaque composante α :

$$\boxed{c_i^\alpha = \sum_{j=1}^p x_{ij} b_j^\alpha} \quad (1), \quad i \in [m], \alpha \in [q], \text{ avec } x_{ij} = \frac{f_{ij}}{f_{i.}} = f_{j/i}.$$

La modalité i est donc représentée au barycentre des modalités j . Ce résultat est visible au départ puisque la modalité j est assimilée à l'individu-variables j dont le profil est : $(0, 0, \dots, 0, 1, 0, \dots, 0)$. On a donc pour chaque

modalité i le profil : $X_i = \sum_{j=1}^p x_{ij}e_j = \sum_{j=1}^p f_{j/i}(0, \dots, 1, \dots, 0)$.

La première expression valable dans toute ACP prend la forme barycentrique de la seconde dès que les x_{ij} , $j \in [p]$ constituent un système de poids, ce qui est le cas dans l'AFC avec $x_{ij} = f_{j/i}$.

c) Comparaison avec les solutions classiques

Les relations barycentriques exactes indiquées ci-dessus ne sont pas celles adoptées dans l'AFC car on tient compte des relations symétriques obtenues dans l'ACP du second triplet qui donne :

$$c_j^{\alpha} = \sum_{j=1}^m y_{ji} b_i^{\alpha} \quad (2), j \in [p], \alpha \in [q], \text{ avec :}$$

$$y_{ji} = \frac{f_{ji}}{f_j} = f_{i/j},$$

c^{α} vecteur propre de YX , D_p -normé à $\sqrt{\lambda_{\alpha}}$,

b^{α} vecteur propre de XY , D_m -normé à 1.

On a donc $c^{\alpha} = \sqrt{\lambda_{\alpha}} b^{\alpha}$ et $c^{\alpha} = \sqrt{\lambda_{\alpha}} b^{\alpha}$, $\alpha \in [q]$. Et comme il n'y a pas plus de raison de prendre (1) ou (2), on choisit de représenter les modalités i et j soit avec c_i^{α} et c_j^{α} , soit avec b_i^{α} et b_j^{α} , soit éventuellement une moyenne des deux. Dans les trois cas on a des composantes homothétiques sur chaque axe Δu_{α} avec des coefficients de proportionnalité égaux à : $\sqrt{\lambda_{\alpha}}$, 1, $\frac{1}{2}(1 + \sqrt{\lambda_{\alpha}})$. C'est habituellement la première écriture qui est retenue avec les formules :

$$c_i^{\alpha} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{j=1}^p \frac{f_{ij}}{f_i} c_j^{\alpha}, \quad i \in [m], \text{ et } c_j^{\alpha} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{i=1}^m \frac{f_{ij}}{f_j} c_i^{\alpha}, \quad j \in [p]; \quad \alpha \in [q].$$

Dans les présentations classiques de l'AFC, ces relations se vérifient à condition de superposer les axes principaux d'inertie des deux ACP ce qui ne reçoit aucune justification précise (Bouroche, Saporta 1980). Ici, au contraire, ayant deux représentations simultanées des modalités $i \in [m]$ et $j \in [p]$, il serait indiqué d'en faire une moyenne comme suggéré ci-dessus en troisième possibilité. On peut aussi adopter la première pour retrouver la présentation traditionnelle.

Les résultats présentés en (1) et (2) sont exactement ceux que donne l'Analyse Canonique des tableaux des indicatrices des modalités des caractères 1 et 2, soient X_1 et X_2 . On verra (par exemple dans l'ouvrage cité ci-dessus) qu'il s'obtiennent au moyen d'une sorte d'artifice : On ne peut pas projeter les variables i et j (ici les indicatrices des modalités) sur le plan des variables canoniques (ξ^1, ξ^2) ou (η^1, η^2) à cause de leur non centrage. On choisit donc de représenter dans l'espace

des individus (et non celui des variables) chaque modalité comme barycentre des individus qui présentent cette modalité. Ce procédé est comparable à celui qui permet de faire figurer les variables de toute ACP dans l'espace des individus : "l'individu" vecteur unitaire e_j de E est assimilable à la variable j et il donne sa direction dans E .

conclusion

Les propriétés de dualité permettent d'orienter l'espace des individus et ses plans principaux d'inertie par l'indication des directions des variables initiales Δe_j , $j \in J$. Dans le cas de l'AFC, on obtient deux représentation simultanées des nuages des modalités, identiques à celles proposées par l'Analyse Canonique pour les indicatrices des modalités des deux caractères. Ainsi les représentations simultanées classiques avec c^α et c'^α se trouvent justifiées. On pourrait aussi utiliser les b^α et b'^α ou n'importe quel barycentre des b^α et c^α pour positionner les modalités i du caractère 1, et le même barycentre des b'^α et c'^α pour positionner les modalités j du caractère 2.

ANNEXE II : L'Intrastructure Discriminante

Nous allons montrer que l'Analyse Canonique Généralisée (ACG) équivaut à la recherche d'une intrastructure par la méthode d'ACP d'un tableau moyen dans \mathbf{R}^p , exactement comme l'AFJ1. Les nuages \tilde{N}_k sont encore \mathcal{M} -projetés sur le sous-espace engendré par les individus du tableau moyen et seule la métrique \mathcal{M} est changée. Ce sous-espace coïncide avec le sous-espace engendré par les axes discriminants de l'Analyse Factorielle Discriminante. Ce dernier point a été montré par Casin (1985) et Turlot, Casin (1986). Ils ont introduit une intrastructure dont nous reprenons l'étude pour en préciser la formulation, les propriétés et les rapports avec les autres méthodes. Nous l'appelons Intrastructure Discriminante.

Après avoir noté avec Saporta (1975) que l'ACG s'obtient par l'ACP du triplet $[X_a = (X_1 \quad \dots \quad X_m), V_{X_a}^{-1}, D]$, Turlot et Casin montrent qu'elle équivaut à l'AFD visant à reconstituer le mieux possible les groupes d'individus de même indice quand on dispose les nuages N_k dans les sous-espaces orthogonaux \mathbf{R}^{p_k} de \mathbf{R}^p . On peut alors définir une intrastructure en projetant les individus des nuages \tilde{N}_k sur le sous-espace \mathbf{R}^s engendré par les s ($s \leq q = \text{rang} X_a$) premiers axes discriminants.

Il est intéressant de préciser ces divers points.

a) L'Analyse Canonique Généralisée.

L'ACG au sens de Carroll recherche les variables z^j , D -orthogonales dans \mathbf{R}^n , qui maximisent successivement $\sum_{k \in K} R^2(z^j, X_k)$, $j \in [q]$, où $R^2(z^j, X_k) = R_{jk}^2$ est le carré du coefficient de corrélation multiple entre z^j et les variables du tableau X_k . On a $R_{jk}^2 = \cos^2 \theta_{jk}$, où θ_{jk} est l'angle entre z^j et $z_k^j = P_k z^j$, P_k désignant le D -projecteur dans $F = (\mathbf{R}^n, D)$ sur le sous-espace F_k engendré par les variables X_k^j , $j \in J_k$ (La notation P_k de cette annexe n'a donc pas la même signification qu'au chapitre III). P_k étant idempotente et D -symétrique, on a : $P_k' D P_k = D P_k$. On en déduit :

$$\sum_{k \in K} R_{jk}^2 = \sum_{k \in K} \frac{z^{j'} D P_k z^j}{z^{j'} D z^j} = \frac{z^{j'} D (\sum_k P_k) z^j}{z^{j'} D z^j}.$$

Les z^j qui maximisent $\sum_k R_{jk}^2$ sont donc les vecteurs propres successifs de la matrice $P = \sum_k P_k = \sum_{k \in K} X_k (X_k' D X_k)^{-1} X_k' D$ associés aux valeurs propres positives décroissantes λ_j avec $\sum_k R_{jk}^2 = \lambda_j$ (on suppose ici : $\text{rang} X_k = p_k$,

$k \in K$). $P = WD = XH^{-1}X'D$ pour l'ACP du triplet (X, H^{-1}, D) avec $X = X_a = (X_1 \quad \dots \quad X_m)$, et $H^{-1} = \text{diag}(V_{XX}^{-1})$ (la notation $\text{diag}(V_{XX})$ désigne la matrice bloc-diagonale contenant en bloc k , la matrice de variances covariances des variables de J_k). Et comme les variables de l'ACG ne sont définies qu'à un facteur près, on peut choisir comme variables canoniques les composantes principales obtenues dans l'ACP du triplet (X, H^{-1}, D) ce qui donne le tableau $(n, q) : Z = C = XH^{-1}U = XB$.

U est le tableau (p, q) des axes principaux d'inertie V_{XX}^{-1} -orthonormés, $B(p, q)$ est celui des facteurs H -orthonormés et V -orthogonaux de normes $\sqrt{\lambda_i}$.

$$Z = XB \text{ s'écrit par blocs : } Z = (X_1 \quad \dots \quad X_m) \begin{pmatrix} B_1 \\ \dots \\ \vdots \\ \dots \\ B_m \end{pmatrix} = \sum_k X_k B_k,$$

soit $Z = C = \sum_k C_k$ avec $C_k = X_k B_k$ (n, p_k).

On montre alors que $z_k^j = P_k z^j = \lambda_j C_k^j = \lambda_j X_k B_k^j$ ou encore $Z_k = C_k \Lambda$, $k \in K$.

On a aussi $Z = \sum_k C_k = \sum_k Z_k \Lambda^{-1}$, ou dans une écriture classique :

$$z^j = \frac{1}{\lambda_j} \sum_{k \in K} z_k^j, \quad j \in [q].$$

On notera que si tous les espaces F_k sont confondus les z_k^j coïncident avec z^j et $\lambda_j = \sum_{k \in K} R_{jk}^2 = m$, $j \in [q]$. Nous qualifierons les z_k^j de composantes canoniques et les z^j de variables canoniques.

b) ACG et Analyse Factorielle Discriminante.

L'ACG donne alors lieu à une interprétation en terme d'ACP à condition de lui associer une Analyse Factorielle Discriminante.

Soit \tilde{X}_e le tableau des $\tilde{X}_k, k \in K$ empilés :

$$\tilde{X}_e = \begin{pmatrix} \tilde{X}_1 \\ \dots \\ \vdots \\ \dots \\ \tilde{X}_m \end{pmatrix} = \begin{pmatrix} \boxed{X_1} & 0 & \dots & 0 \\ 0 & \boxed{X_2} & \dots & 0 \\ 0 & & \ddots & 0 \\ 0 & & \dots & \boxed{X_m} \end{pmatrix}.$$

Il réunit l'ensemble des nuages $\tilde{N}_k, k \in K$ positionnés dans les sous-espaces supplémentaires orthogonaux de $(\mathbf{R}^p, \mathcal{M})$. Pour reconstituer au mieux les n classes formés par les individus de même indice : $\{x_i^k, k \in K\}$ parmi les mn individus, l'AFD recherche un axe $\Delta\omega$ de \mathbf{R}^p tel que pour l'ensemble des mn projections sur

cet axe, on maximise la dispersion entre groupes (l'inertie inter B_ω) et minimise la dispersion à l'intérieur des groupes (l'inertie intra W_ω).

Soient la métrique \mathcal{M} dans \mathbf{R}^p , ω le vecteur unitaire de $\Delta\omega$, $b = \mathcal{M}\omega$ le facteur associé ; T, W et R les matrices de variances-covariances totale, intra et inter (le symbole R remplace l'habituel B qui désignera ici la matrice des facteurs). En supposant T et W inversibles, le problème s'écrit :

$$\begin{aligned} \max_{\omega} \frac{B_\omega}{W_\omega} &\iff \max_b \frac{b'Rb}{b'Wb} \iff W^{-1}Rb = \mu b \\ &\iff W^{-1}Tb = (\mu + 1)b \\ &\iff \max_b b'Rb \text{ sous } b'Tb = a \\ &\iff T^{-1}Rb = \lambda b \quad \left(\lambda = \frac{\mu}{1 + \mu}\right) \end{aligned}$$

Il existe donc trois ACP donnant les axes discriminants successifs :

- les deux ACP du nuage des centres de gravité, donc du tableau compromis $X_\rho = \sum_k \rho_k \tilde{X}_k$ avec les métriques W^{-1} ou T^{-1} ,
- l'ACP du nuage de tous les points, donc du tableau $\tilde{X}_e(mn, p)$ avec la métrique W^{-1} .

L'égalité $T = R + W$ avec $T = \tilde{X}'D_\rho\tilde{X}$ et $R = X'_\rho DX_\rho$ s'écrit :

$$\begin{aligned} \begin{pmatrix} \rho_1 V_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_m V_{mm} \end{pmatrix} &= \begin{pmatrix} \rho_1^2 V_{11} & \cdots & \rho_1 \rho_m V_{1m} \\ \vdots & \ddots & \vdots \\ \rho_m \rho_1 V_{m1} & \cdots & \rho_m^2 V_{mm} \end{pmatrix} \\ &+ \begin{pmatrix} \rho_1(1 - \rho_1)V_{11} & \cdots & -\rho_1 \rho_m V_{1m} \\ \vdots & \ddots & \vdots \\ -\rho_m \rho_1 V_{m1} & \cdots & \rho_m(1 - \rho_m)V_{mm} \end{pmatrix} \end{aligned}$$

L'AFD qui s'obtient par l'ACP du triplet (X_ρ, T^{-1}, D) est alors, dans l'hypothèse où les poids des tableaux sont égaux ($\rho_k = \frac{1}{m}, k \in K$), identique (à un facteur près) à l'ACG, qui est, l'ACP du triplet (X, H^{-1}, D) .

c) L'Intrastructure Discriminante.

Comparons les résultats des ACP de ces deux triplets en notant les éléments calculés pour le triplet discriminant (X_ρ, T^{-1}, D) , avec les symboles relatifs à l'ACP du triplet (X, H^{-1}, D) indicés par d .

On a : $X_\rho = \frac{1}{m}X$ et $T^{-1} = mH^{-1} \implies W_d = X_\rho T^{-1} X'_\rho = \frac{1}{m}W$, $V_d = \frac{1}{m^2}V$,
 $W_d D = \frac{1}{m}WD$ et $T^{-1}V_d = \frac{1}{m}H^{-1}V \implies \lambda_d = \frac{1}{m}\lambda$ (pour tout λ_i). Les
facteurs discriminants b (notés b_d) sont les vecteurs propres T -orthonormés de
 $T^{-1}V_d = \frac{1}{m}H^{-1}V$ et $b'_j T b_j = 1 \iff b'_d \frac{H}{m} b_d = 1 \implies b_d = \sqrt{m}b$ ou $B_d = \sqrt{m}B$
de même pour les composantes principales :

$$C'_d D C_d = \frac{1}{m} \Lambda \implies C_d = \frac{1}{\sqrt{m}} C.$$

Cette relation entre variables canoniques et position des individus compromis dans la base discriminante se double de relations entre composantes canoniques et intrastructure dans cette base.

En effet soit $B_d = \sqrt{m}B$ la matrice des facteurs telle que $C_d = X_\rho B_d$. Elle permet de calculer les projections des nuages définis par $\tilde{X}_k, k \in K$ sur la base discriminante avec $\hat{X}_k = \tilde{X}_k B_d, k \in K$, soit :

$$\hat{X}_k = (0 \quad \dots \quad X_k \quad \dots \quad 0) \sqrt{m} \begin{pmatrix} B_1 \\ \dots \\ \vdots \\ \dots \\ B_m \end{pmatrix} = \sqrt{m} X_k B_k = \sqrt{m} Z_k \Lambda^{-1}.$$

L'AFD ainsi exposée, n'est rien d'autre qu'une méthode d'ACP du tableau moyen dans \mathbf{R}^p avec la métrique T^{-1} , exactement comme l'AFJ1 qui utilise la métrique \mathcal{M} .

Aussi avec le même argument que dans l'AFJ1 nous introduisons un facteur \sqrt{m} dans l'intrastructure qui devient : $\hat{X} = C$; $\hat{X}_k = m Z_k \Lambda^{-1}, k \in K$.

Résumé des propriétés de l'ACG et l'AFD

- En plus de leurs propriétés spécifiques, les variables canoniques de l'ACG au sens de Carroll définissent une intrastructure des nuages \tilde{N}_k positionnés dans des sous-espaces supplémentaires de \mathbf{R}^p :

- les variables canoniques, assimilées aux colonnes du tableau C, donnent le positionnement des individus compromis (pour des poids égaux : $\rho_k = \frac{1}{m}$),
- les composantes canoniques Z_k^j relatives au tableau X_k fournissent, au facteur $\frac{m}{\lambda_j}$ près, les coordonnées des n individus du tableau X_k sur les axes $\Delta\omega_j, j \in [q]$, de l'intrastructure.

- Sachant que les composantes canoniques sont déterminées à un facteur près, si on définit les composantes canoniques z_k^j comme les projections de z^j sur F_k

multipliées par le facteur $\frac{m}{\lambda_j}$, alors les composantes canoniques donnent exactement l'intrastructure recherchée. C'est l'Intrastructure Discriminante.

• L'ACG est donc identique au problème de la recherche d'une Intrastructure Discriminante pour laquelle les individus de même indice sont regroupés le mieux possible autour de leur centre de gravité.

d) Particularités de l'Intrastructure Discriminante.

L'AFD contient cependant une subtilité non présente dans l'AFJ1. Elle est bien méthode du tableau moyen dans \mathbf{R}^p muni de la métrique T^{-1} au sens où son intrastructure s'obtient par l'ACP du tableau moyen avec T^{-1} -projection des nuages \tilde{N}_k sur les axes principaux du compromis. Mais elle s'interprète dans l'espace $E^p = \mathbf{R}^p$ muni de sa métrique \mathcal{M} comme projection du nuage compromis et des nuages \tilde{N}_k sur les axes discriminants $\Delta\omega_j$, $j \in [q]$.

d1) Qualité de la représentation du compromis.

L'AFJ1 (resp. l'ACG) s'obtient par l'ACP du tableau moyen X_ρ avec la métrique \mathcal{M} (resp. T^{-1}) et les composantes principales donnent les coordonnées des individus compromis dans une base du sous-espace E_q de \mathbf{R}^p engendré par les individus compromis.

Pour l'AFJ1 cette base U est \mathcal{M} -orthonormée et les c^j sont les \mathcal{M} -projections des n individus compromis sur les vecteurs u_j de la base, $j \in [q]$.

Pour l'ACG et l'AFD cette base U est T^{-1} -orthonormée et les c^j sont encore les T^{-1} -projections mais ce sont aussi des \mathcal{M} -projections sur les vecteurs ω_j , $j \in [q]$ formant une base Ω du sous-espace discriminant.

Proposition 1.- *Dans l'AFJ1, la représentation du compromis dans l'intrastructure se fait avec une i.e.e. (image euclidienne équivalente) à celle de départ.*

De plus les premières composantes principales restituent le maximum d'inertie de cette image.

■ Notons \tilde{C}_I ($n \times p$) la matrice de composantes principales C_I de l'AFJ1 complétée avec des colonnes de 0, \tilde{U} la base U complétée de vecteurs \mathcal{M} -orthonormés pour former une base de \mathbf{R}^p . On a : $\tilde{U}'\mathcal{M}\tilde{U} = I_p \implies \tilde{U}'\mathcal{M} =$

$\tilde{U}^{-1} \implies \tilde{U}\tilde{U}'\mathcal{M} = I_p$. Avec une écriture par blocs, on en déduit : $UU'\mathcal{M} = I_p = \mathcal{M}UU'$ et $\tilde{C}_I I_p \tilde{C}_I' = C_I I_q C_I'$. Alors :

$$\tilde{C}_I = X_\rho \mathcal{M} \tilde{U} \implies \tilde{C}_I I \tilde{C}_I' = X_\rho \mathcal{M} \tilde{U} \tilde{U}' \mathcal{M} X_\rho' = X_\rho \mathcal{M} X_\rho' = C_I I C_I' \quad \blacksquare$$

Proposition 2.- *Dans l'ACG, la représentation du compromis dans l'intrastructure définit une i.e. non équivalente à celle du départ, car elle se lit avec la métrique I . On aurait une i.e.e. en prenant la métrique $\mathcal{N} = U'\mathcal{M}U$. La base Ω , nommée Repère Discriminant, est $\mathcal{M}T\mathcal{M}$ -orthonormale.*

■ • \tilde{U} étant une base T^{-1} orthonormée de \mathbf{R}^p on a :

$$\tilde{U}'T^{-1}\tilde{U} = 1 \implies T^{-1}UU' = I_p = UU'T^{-1}$$

$$C = X_\rho T^{-1}U \implies CNC' = X_\rho (T^{-1}U)U'\mathcal{M}U(U'T^{-1})X_\rho' = W.$$

• Il y a donc équivalence des triplets : (X_ρ, \mathcal{M}, D) et (C, \mathcal{N}, D) avec $\mathcal{N} = U'\mathcal{M}U$ et donc pas d'équivalence — dans le cas général — des triplets (X_ρ, \mathcal{M}, D) et (C, I, D) . Cette dernière équivalence se vérifie si $\mathcal{M} = T^{-1}$ et dans ce cas l'AFJ1, l'ACG et l'AFD sont équivalentes (par exemple si les variables de chaque tableau sont D -orthogonales, avec la métrique de l'ACP normée, on a : $\mathcal{M} = T^{-1}$).

• On a les mêmes facteurs pour définir les T^{-1} -projections sur U ou les \mathcal{M} -projections sur Ω : $T^{-1}U = \mathcal{M}\Omega = B$. De $B'TB = I$ on tire $\Omega'\mathcal{M}T\mathcal{M}\Omega = I$ et la base Ω est $\mathcal{M}T\mathcal{M}$ -orthonormale. Elle est aussi \mathcal{M} -orthonormale par définition. ■

d2) Qualité de l'intrastructure.

Comme l'AFD est une AFJ1 pour la métrique T^{-1} , on est sûr d'avoir une intrastructure inexacte, sauf pour des tableaux égaux, relativement à la métrique T^{-1} . Cette propriété se vérifie encore dans l'espace $E^p = (\mathbf{R}^p, \mathcal{M})$. Nous nous contenterons de le vérifier dans un cas particulier.

Proposition 3.- *L'Intrastructure Discriminante de tableaux proportionnels fait coïncider tous les nuages avec le nuage compromis : $\hat{X} = \hat{X}_k = C$, $k \in K$.*

■ On a $X_k = a_k X$ et $M_k = V_k^{-1}$, $k \in K$, avec $\sum_k a_k^2 = m$. Soient $V = X'DX$ et $W = XV^{-1}X'$. On a : $V_k = a_k^2 V$ et $W_k = a_k^2 X a_k^{-2} V^{-1} X' = W$, $k \in K$. Tous les nuages sont confondus dans l'intrastructure puisque les W_k , $k \in K$ sont égaux. ■

Cette proposition montre qu'à l'inverse de l'intrastructure de l'AFJ1, cette intrastructure dilate les petits nuages et réduit les gros.

Elle illustre bien la propriété de l'Intrastructure Discriminante de regrouper "le mieux possible" les individus \hat{x}_i^k autour de leur compromis \hat{x}_i sur les premiers axes. En résumé :

Proposition 4.-

Pour une famille de triplets indexés (X_k, M_k, D) , $k \in K$, on a :

- 1) L'ACP du triplet $(X_a = X, V_{XX}^{-1}, D)$ qui donne avec les \sqrt{m} -projections des \tilde{X}_k en supplémentaires, deux résultats indépendants des métriques M_k :
 - L'Intrastructure Discriminante $\hat{X} = C$ et $\hat{X}_k = W_k D C A^{-1}$, avec $W_k = X_k V_k^{-1} X_k'$, $k \in K$.
 - L'Analyse Canonique Généralisée avec :
 - Les variables canoniques : $Z = C$,
 - les composantes canoniques : $Z_k = P_k Z = \frac{1}{m} \hat{X}_k \Lambda = \frac{1}{m} W_k D C$, $k \in K$.
- 2) Les équivalences entre les énoncés suivants :
 - i) Les sous-espaces F_k ont s directions communes Δy^j , $j \in [s]$ dans \mathbf{R}^n .
 - ii) Les s premières composantes canoniques sont égales pour les différents tableaux : $Z_k^s = Z^s$, $k \in K$.
 - iii) $\forall j \in [s] : \lambda_j = m$.
 - iv) Sur les s premiers axes $\Delta \omega_j$ de l'Intrastructure Discriminante, tous les individus x_i^k ont la même coordonnée : $\forall j \in [s], \forall i \in I, \forall k \in K : \hat{x}_i^{jk} = \hat{x}_i^j$.

Remarques :

— Les axes $\Delta \omega_j$ ne sont pas nécessairement des directions de grande inertie du nuage compromis : les $\Delta \omega_j$ le sont pour la métrique V_{XX}^{-1} , les seconds pour la métrique \mathcal{M} .

— La base $(\omega_j)_{j \in [q]}$ est \mathcal{M} et $MV_{XX}\mathcal{M}$ -orthonormale. C'est le seul endroit où la métrique \mathcal{M} intervient.

— Si l'AFJ paraît inapplicable, cette méthode montre avec l'ACG que des variables de grande norme dans F sont loin de la proportionnalité, et avec l'Intrastructure Discriminante, les individus éventuellement responsables de cette situation. En supprimant ces variables ou individus, on peut retrouver de bonnes conditions d'application de l'AFJ.

ANNEXE III : Analyse Procuste et AFJ

Proposition - Si la famille de triplets vérifie : $C_k^q = CA_k$, $k \in K$ alors l'Intrastructure Equisecteur réalise l'Analyse Procuste Généralisée des tableaux \tilde{X}_k . Dans ce cas celle-ci acquiert la signification statistique de l'AFJ en plus de sa signification géométrique habituelle.

■ Rappelons que le problème procuste généralisé pour m triplets (X_k, M, D) , $k \in K$ avec $X_k(n, \bar{p})$ et un système de poids ρ_k , est de trouver des isométries H_1, H_2, \dots, H_m de $E^{\bar{p}} = (\mathbf{R}^{\bar{p}}, M)$ dans lui-même réalisant le minimum de :

$$\sum_{k \in K} \sum_{\ell \in K} \rho_k \rho_\ell \|X_k H_k - X_\ell H_\ell\|_{\varphi_{DM}}^2 .$$

$$\forall (k, \ell) \in K^2 \text{ , on a : } \|X_k H_k - X_\ell H_\ell\|_{\varphi_{DM}}^2 = \|(X_k - X_\ell H_\ell H_k^{-1}) H_k\|_{\varphi_{DM}}^2 \\ = \|X_k - X_\ell H_\ell H_k^{-1}\|_{\varphi_{DM}}^2 .$$

Prenons d'abord les tableaux $C_k^q(n, q)$ et l'espace $E_q = (\mathbf{R}^q, I_q)$. Pour minimiser $S = \sum_k \sum_\ell \rho_k \rho_\ell \|C_k^q H_k - C_\ell^q H_\ell\|_{\varphi_{ID}}^2$, nous commençons par minimiser séparément chaque terme de la somme. Cela consiste à résoudre l'analyse procuste des 2 tableaux C_k^q, C_ℓ^q . On cherche donc l'isométrie $H_{k\ell} = H_\ell H_k^{-1}$ réalisant le minimum de :

$$\|C_k^q - C_\ell^q H_{k\ell}\|_{\varphi_{ID}}^2 = \|C_k^q\|_{\varphi_{ID}}^2 + \|C_\ell^q H_{k\ell}\|_{\varphi_{ID}}^2 - 2 \text{trace}(C_k^{q'} D C_\ell H_{k\ell}) .$$

Le problème revient à maximiser $\text{trace}(C_k^{q'} D C_\ell H_{k\ell})$. La s.v.d. permet d'écrire $C_k^{q'} D C_\ell^q = V \Lambda U'$ avec $U(q, q), V(q, q)$ orthogonales et Λ diagonale s.d.p. . Comme $C_k^{q'} D C_\ell^q = \text{diag}(a_{kj} a_{\ell j} \|c^j\|^2)$, avec $a_{kj} a_{\ell j} \geq 0$, on peut choisir $U = V = I_q$. Alors $H_{k\ell} = UV' = I_q$ est une solution cherchée. Elle est unique si $\text{rang} C_k^q = q_k = q_\ell = q$ (Glaçon 1981). Les solutions $H_k = H_\ell = I_q$ étant identiques pour tous les couples (k, ℓ) , on a bien la solution du problème procuste général (symétrisé).

On a donc :

$$\sum_k \sum_\ell \rho_k \rho_\ell \|C_k^q - C_\ell^q\|_{\varphi_{ID}}^2 = \text{Min}_{\hat{H}_k, \hat{H}_\ell} \sum_k \sum_\ell \rho_k \rho_\ell \|C_k^q H_k - C_\ell^q H_\ell\|_{\varphi_{ID}}^2 .$$

Comme dans l'Intrastructure Equisecteur on a $\hat{C}_k = C_k^q$, $k \in K$ et que les triplets (C_k^q, I_q, D) et $(\tilde{X}_k, \mathcal{M}, D)$ sont équivalents on a aussi le résultat annoncé :

$$\sum_k \sum_\ell \rho_k \rho_\ell \|\hat{C}_k - \hat{C}_\ell\|_{\varphi_{ID}}^2 = \text{Min}_{\tilde{H}_k, \tilde{H}_\ell} \sum_k \sum_\ell \rho_k \rho_\ell \|\tilde{X}_k \tilde{H}_k - \tilde{X}_\ell \tilde{H}_\ell\|_{\varphi_{DM}}^2 .$$

Sous l'hypothèse considérée, AFJ et analyse procustéenne coïncident et l'Intra-structure obtenue possède les propriétés liées aux deux méthodes :

- la représentation simultanée respecte la règle de dualité pour laquelle chaque axe du sous-espace Equisecteur E_E est associée à une seule direction dans l'espace F des variables.
- la somme des carrés des distances entre individus de même indice est minimale.

ANNEXE IV : Comparaison des sous-espaces Compromis et Déformant

Sous l'hypothèse de composantes proportionnelles : $Y_k = YA_k$, $k \in K$, la proposition 6 établit la coïncidence des sous-espaces Compromis E_C et Déformant E_D . Donc l'initialisation de l'algorithme de l'AFJ par le tableau compromis C est optimale car il coïncide avec le tableau C_Y des composantes principales de Y . Nous allons montrer que cette initialisation reste excellente pour les tableaux à composantes quasi-proportionnelles en distinguant deux cas. Dans le premier cas, on montre une très grande proximité de E_C et de E_D , donc une très grande ressemblance des deux tableaux C et C_Y précédents. Dans le second cas, on verra E_C s'écarter légèrement de E_D . Alors l'Intrastructure Compromis obtenue en projection sur E_C peut différer très largement de l'Intrastructure Déformante obtenue en projection sur E_D . Cela n'empêche pas les tableaux C et C_Y de rester suffisamment ressemblants pour que l'initialisation proposée soit excellente.

Supposons que les tableaux soient à composantes quasi-proportionnelles. Pour simplifier les écritures, nous prenons les poids ρ_k égaux et $E^{p_c} = E^{mq}$. Considérons $E^{mq} = \bigoplus E_k^q$ rapporté à la base $(\tilde{v}_j^k)_{j \in [q], k \in K}$ qui permet de définir le R.D. : $\{v_j = \sqrt{m} \sum_{k \in K} \rho_k a_{kj} \tilde{v}_j^k, j \in [q]\}$. Dans E_k^q on a $Y_k = YA_k + \varepsilon_k$, $k \in K$, d'où le vecteur \tilde{y}_i^k donnant les coordonnées de l'individu i du tableau k dans la base (\tilde{v}_j^k) :

$$\tilde{y}_i^k = \sum_{j \in [q]} y_i^{jk} \tilde{v}_j^k = \sum_{j \in [q]} a_{kj} y_i^j \tilde{v}_j^k + \sum_{j \in [q]} \varepsilon_i^{jk} \tilde{v}_j^k, \quad i \in I, k \in K.$$

L'individu compromis est le \sqrt{m} -barycentre des \tilde{y}_i^k , $k \in K$, et $\forall i \in I$:

$$\bar{y}_i = \sqrt{m} \sum_k \rho_k \tilde{y}_i^k = \sqrt{m} \sum_k \frac{1}{m} \sum_j a_{kj} y_i^j \tilde{v}_j^k + \frac{1}{\sqrt{m}} \sum_k \sum_j \varepsilon_i^{jk} \tilde{v}_j^k,$$

$$\bar{y}_i = \sum_j y_i^j \frac{1}{\sqrt{m}} \sum_k \tilde{v}_j^k + \frac{1}{\sqrt{m}} \sum_k \sum_j \varepsilon_i^{jk} \tilde{v}_j^k = \sum_j y_i^j v_j + \frac{1}{\sqrt{m}} \sum_k \sum_j \varepsilon_i^{jk} \tilde{v}_j^k. \quad (1)$$

Si les tableaux sont à composantes proportionnelles, on retrouve évidemment le résultat : $\bar{Y} = Y$. Cette expression va surtout nous permettre de montrer que, selon la distribution des ε_i^{jk} , les individus compromis \bar{y}_i peuvent être plus ou moins proches de leur homologue de l'Intrastructure Déformante : $y_i = \sum_j y_i^j v_j$, $i \in I$. On comprend ainsi pourquoi l'I.C. peut être très proche ou très éloignée de l'I.D.. Cette différence tient à la distribution des ε_i^{jk} , alors que la différence entre I.D. et I.E. tient à l'écart des coefficients a_{kj} à l'unité.

a) Coïncidence du R.C. et du R.D.

L'égalité (1) s'écrit : $\bar{y}_i = y_i + \sum_k \sum_j \frac{1}{\sqrt{m}} \varepsilon_i^{jk} \tilde{v}_j^k$, $i \in I$, et on peut l'exprimer dans une base de E^{mq} obtenue en complétant la base (v_j) du sous espace Déformant E_D par $(m-1)q$ vecteurs orthonormaux notés : v_{kj} , $k \in [m] - \{1\}$, $j \in [q]$. On a ainsi :

$$\bar{y}_i = \sum_j (y_i^j + \alpha_i^{j1}) v_j + \sum_{k=2}^m \sum_j \alpha_i^{jk} v_{kj}, \quad i \in I.$$

$\forall (i, j, k) \in I \times [q] \times K$, supposons les $|\varepsilon_i^{jk}|$ suffisamment petits relativement à $|y_i^j|$ pour vérifier la relation : $|\varepsilon_i^{jk}| < \iota |y_i^j|$ (par exemple, $\iota = 10\%$). Dans une isométrie, le vecteur $\bar{y}_i - y_i$ conserve sa norme donc $\sum_k \sum_j \frac{1}{m} (\varepsilon_i^{jk})^2 = \sum_k \sum_j (\alpha_i^{jk})^2$. En supposant les $|\alpha_i^{jk}|$ approximativement égaux entre eux, ainsi que les $|\varepsilon_i^{jk}|$, on obtient : $|\alpha_i^{jk}| \simeq \frac{1}{\sqrt{m}} |\varepsilon_i^{jk}| < \frac{\iota}{\sqrt{m}} |y_i^j|$ (Cette dernière hypothèse très grossière nous sert simplement à imaginer un des cas possibles). Les composantes $y_i^j + \alpha_i^{j1}$ de \bar{y}_i sur (v_j) diffèrent fort peu de celle de y_i . Par ailleurs les \bar{y}_i admettent des composantes sur (v_{kj}) de très faible valeur absolue, et celles de y_i sont nulles. Comparons maintenant les axes principaux d'inertie des nuages : $\{\bar{y}_i, i \in I\}$ et $\{y_i, i \in I\}$. Ces points ont pratiquement les mêmes q premières composantes et les \bar{y}_i ont en plus $(m-1)q$ composantes très petites, ne donnant que très peu d'inertie au nuage compromis dans les directions de E^{p_c} n'appartenant pas au sous-espace déformant E_D .

On voit donc la possibilité d'avoir les q premiers axes d'inertie des deux nuages pratiquement confondus : il y a quasi coïncidence du R.D.P. avec les q premiers axes du R.C. (mais ce dernier admet d'autres vecteurs). Il en résulte que I.D.P. et I.C. sont quasi confondues c'est à dire d'autant plus ressemblantes que les $|\varepsilon_i^{jk}|$ sont petits et vérifient les hypothèses introduites.

L'exemple des "cantons profils" du chapitre IV relève de ce cas. Il est important de noter que la ressemblance de l'I.C. avec l'I.D.P. laisse entrevoir des cas où l'I.C. de l'AFJ1 (Statis, AFM) peut être presque exacte donnant ainsi un champ d'application à ces méthodes. Il suffit que les conditions précédentes soient remplies (coïncidence de l'I.C. et de l'I.D.P) et que les coefficients a_{kj} vérifient : $A_k \simeq I_q$, $k \in K$ (coïncidence de l'I.D.P. et de l'I.E.P.). La condition : $Y_k = Y + \varepsilon_k$ avec des $|\varepsilon_i^{jk}|$ suffisamment petits et égaux entre eux, est donc suffisante pour avoir une I.C. exacte car elle est alors quasi identique à l'I.E.P..

Nous allons maintenant examiner la situation parfaitement opposée et voir surgir de grosses différences entre I.C. et I.D.P..

b) L'effet P.L.M.

Nous restons dans le cas des composantes quasi proportionnelles. Nous supposons encore : $|\varepsilon_i^{jk}| < \iota |y_i^j|$, sauf pour un individu $n \in I$, pour lequel on a :

- $\exists l \in K, |\varepsilon_n^{ll}| > \iota |y_n^l|$ (par exemple : $\varepsilon_n^{ll} = 0.4 \text{Max}_i |y_i^l|$, $\iota = 10\%$), et

- $\exists j > 1, |y_n^j| = \text{Max}_i |y_i^j|$.

Supposons de plus que le nuage compromis est "long" avec $\|y^1\| \gg \|y^2\| \gg \|y^3\| \gg \dots$ de telle sorte que seuls les $|\varepsilon_i^{1k}|$ relatifs à la composantes y^1 ont une valeur non négligeable. Avant d'illustrer ces hypothèses sur un exemple, rappelons que $\sum_k \varepsilon_k A_k = 0$ (Proposition 10), et montrons que :

Lemme. Dans l'écriture $Y_k = Y A_k + \varepsilon_k$ les variables Y_k^j, Y^j et ε_k^j sont centrées donc : $\sum_{i \in I} p_i \varepsilon_i^{jk} = 0, j \in [q], k \in K$.

■ On a $Y_k = C_k Q_k'$ et $Y = C Q_k'$ avec les variables C_k et C centrées (ce sont des composantes principales). Le vecteur de F ayant toutes ses composantes égales à 1 étant noté $\mathbf{1}_n$, on a : $C' D \mathbf{1}_n = 0 \Rightarrow Y' D \mathbf{1}_n = Q C' D \mathbf{1}_n = 0$. Et de même pour Y_k . Alors : $\varepsilon_k' D \mathbf{1}_n = (Y_k' - A_k Y') D \mathbf{1}_n = 0$. ■

exemple

Une famille de 4 tableaux vérifie $Y_k = Y A_k$ (et même $C_k = C A_k$ sur la figure qui suit) sauf pour les individus 1 et 2 tels que :

- $y_i^{1k} = a_{k1} y_i^1 + \varepsilon_i^{1k}, i = 1, 2$ et $k = 1, 4$.

- $y_1^1 = y_2^1 = \text{Max}_i |y_i^1| = 25$.

- $y_1^2 = 15 = \text{Max}_i |y_i^2|$ et $y_2^2 = 2$.

En prenant $a_{11} = a_{41} = 1, \varepsilon_1^{11} = -\varepsilon_1^{14} = -10, \varepsilon_1^{12} = \varepsilon_1^{13} = 0$, on a bien $\sum_k \varepsilon_k A_k = 0$. Pour vérifier les conditions précédentes, on choisit pour l'individu 2 des ε_2^{1k} opposés aux ε_1^{1k} , soit : $\varepsilon_2^{11} = -\varepsilon_2^{14} = 10, \varepsilon_2^{12} = \varepsilon_2^{13} = 0$. L'égalité (1) appliquée aux individus 1 et 2 s'écrit :

$$\bar{y}_1 = y_1 + \sum_k \sum_j \frac{1}{\sqrt{m}} \varepsilon_1^{jk} \bar{v}_j^k = y_1 + \frac{1}{2} (-10 \bar{v}_1^1 + 10 \bar{v}_1^4) = y_1 - 5 \bar{v}_1^1 + 5 \bar{v}_1^4,$$

$$\bar{y}_2 = y_2 + 5 \bar{v}_1^1 - 5 \bar{v}_1^4.$$

La figure 1 compare d'une part, les trajectoires théorique ($Y_{ki} = Y_i A_k$) et effective ($Y_{ki} = Y_i A_k + \varepsilon_{ki}$) des points 1 et 2, d'autre part, les positions des points \bar{y}_i et $y_i, i = 1, 2$.

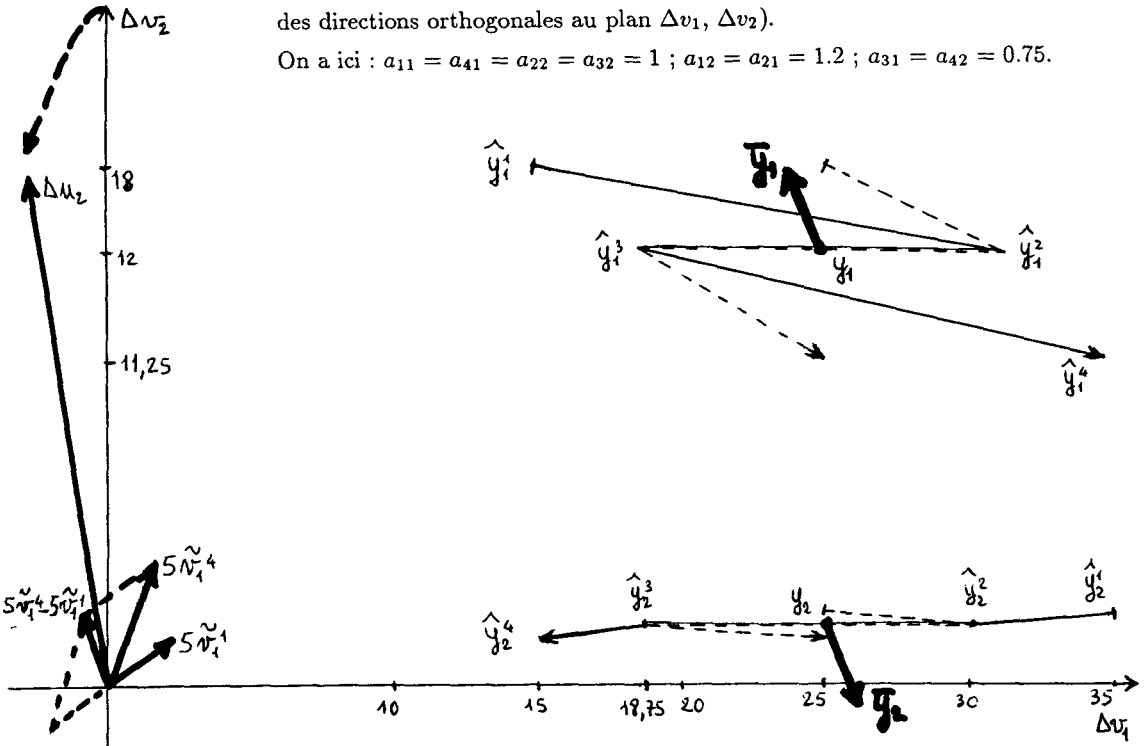
Figure 1 : Position de l'axe Δu_2 du Repère Compromis relativement à l'axe Δv_2 du Repère Déformant.

\dashrightarrow trajectoire théorique ($y_i^{jk} = a_{kj}y_i^j$).

\longrightarrow trajectoire réelle.

\longrightarrow déplacements des points compromis \bar{y}_i de l'I.C. et de l'axe Δu_2 relativement à leurs homologues de l'I.D. (ces déplacements se font essentiellement dans des directions orthogonales au plan $\Delta v_1, \Delta v_2$).

On a ici : $a_{11} = a_{41} = a_{22} = a_{32} = 1$; $a_{12} = a_{21} = 1.2$; $a_{31} = a_{42} = 0.75$.



Sur cet exemple qui caricature l'exemple 1 du chapitre IV, le déplacement des \bar{y}_i relativement aux y_i ($i = 1, 2$) peut avoir les effets suivants :

- stabilité de l'axe 1 du fait que le déplacement $y_1 \rightarrow \bar{y}_1$ est exactement l'opposé du déplacement $y_2 \rightarrow \bar{y}_2$ et que cet axe porte une grande inertie, avec une contribution des points 1 et 2 assez petite (moins de 20 %).
- possibilité de déplacements importants de l'axe 2 surtout si le déplacement du point \bar{y}_2 le laisse près de l'axe 1 alors que le point \bar{y}_1 en est très éloigné et contribue ainsi fortement à l'inertie de l'axe 2 (60% par exemple). Dans ce cas, le point \bar{y}_1 entraîne l'axe 2 dans les directions vers lesquelles il se déplace relativement à y_1 .

L'axe 2 subit ainsi un déplacement vers les axes $\Delta\tilde{v}_1^1$ et $\Delta\tilde{v}_1^4$ d'autant plus sensible que m est grand car cela augmente l'angle θ_j entre chaque axe $\Delta\tilde{v}_j^k$ et l'axe Δv_j associé ($\cos \theta_j = \frac{1}{\sqrt{m}} a_{kj}$ et si tous ces axes sont proches de l'orthogonalité, les projections sur Δv_j sont instables). Donc tous les points ayant de grandes composantes sur ces axes vont voir leur projection sur l'axe 2 contenir d'une part, la composante y_i^2 du point y_i , d'autre part, une projection de la composante y_i^{1k} comptée négativement si $k = 1$ ou positivement si $k = 4$. L'exemple des accidents (chapitre IV) montre parfaitement comment les composantes sur l'axe 2 et surtout l'axe 3, contiennent une partie proportionnelle à la composante 1 du point considéré et nous précisons alors quels ε_i^{jk} vérifient les hypothèses requises.

En revanche, cette disposition particulière des ε_i^{jk} modifie très peu le tableau compromis C qui sert à initialiser l'algorithme de l'AFJ. En effet, seuls les points y_1 et y_2 ont été déplacés. Cela change légèrement la position de l'axe 2 (ou 3, ...) dans l'espace E^{p_c} mais très peu les coordonnées des points du nuage compromis dans le système de ses axes principaux d'inertie.

Enfin, le cas général peut éventuellement être considéré comme combinaison linéaire des deux cas précédents et la conclusion sera la même : sous l'hypothèse de composantes quasi-proportionnelles, le tableau compromis C initialise l'algorithme de l'AFJ le mieux possible du fait de sa stabilité même s'il n'y a pas exacte coïncidence des sous-espaces E_C et E_D .

ABREVIATIONS ET NOTATIONS UTILISEES

1.- Abréviations.

a) Les méthodes

ACG	Analyse Canonique Généralisée
ACP	Analyse en Composantes Principales
AFC	Analyse Factorielle des Correspondances
AFD	Analyse Factorielle Discriminante
AFJ	Analyse Factorielle conJointe
AFJ1	Analyse Factorielle conJointe 1 ^{ère} étape (cf AFM)
AFJ2	2 ^{ème} version de l'AFJ1 (cf Statis)
AFM	Analyse Factorielle Multiple
Candecomp	CANonical DECOMPosition
IDIOSCAL	Individual Differences In Orientation SCALing
INDSCAL	INDividual Differences in muldimensional SCALing
Parafac	PARAllel FACtor analysis
STATIS	Structuration des Tableaux A Trois Indices de la Statistique
s.v.d	singular value decomposition

b) Les objets

I.C.	Intrastructure Compromis
I.D.	Intrastructure Déformante
I.E.	Intrastructure Equisecteur
R.C.(P.)	Repère Compromis (Principal)
R.D.(P.)	Repère Déformant (Principal)
R.E.(P.)	Repère Equisecteur (Principal)
(s.)d.p.	(semi) définie positive
TCE(C)	Tableau Conjoint Empilé multicentré (par défaut ou Centré)
TCA	Tableau Conjoint Aligné

2.- Notations.

a) Pour un tableau unique

I, J	ensembles des n individus et des p variables étudiés
X ou X_{IJ}	tableau ($n \times p$) des observations
X'	transposée de X

$E = (\mathbf{R}^p, M)$ espace métrique des individus, de dimension p , de métrique M
 $F = (\mathbf{R}^n, D)$ espace métrique des variables
 D, p_i métrique sur F assimilée à la matrice diagonale des poids p_i des individus. $D = \text{diag}(p_i)$
 $M, (m_j)$ métrique sur E assimilée à une matrice symétrique d.p.
 Si elle est diagonale, les m_j sont les "poids" des variables.
 I_n ou I matrice identité ($n \times n$)
 X_i, X^j ligne i , colonne j du tableau X
 $x_i = (X_i)'$ individu i considéré comme vecteur de E
 $x^j = X^j$ variable j considérée comme vecteur de F
 x_i^j j -ième coordonnée de x_i ou i -ième coordonnée de x^j
 $\boxed{X'}$ ou N nuage de E défini par les colonnes du tableau X'
 $(e_j)_{j \in J}$ base canonique de E
 Δe_j axe de E engendré par e_j
 $(f^i)_{i \in I}$ base canonique de F
 $\mathbf{R}_{(e_j)}^p$ l'espace \mathbf{R}^p rapporté à la base $(e_j)_{j \in J}$
 $W = XM X'$ matrice ($n \times n$) des produits scalaires entre individus
 $V = X' D X$ matrice ($p \times p$) des variances covariances
 q le rang de X, V ou W
 $\lambda_j, \Lambda = \Delta^2$ valeur propre de WD , matrice diagonale ($q \times q$) des $\lambda_j > 0$ décroissantes
 E^q, F^q sous espaces engendrés par les individus dans E , les variables dans F
 $(u_j)_{j \in [q]}$ vecteurs propres M -orthonormés de VM associés aux q premières valeurs propres. C'est une base orthonormale de E^q, I_q
 $(\gamma_j)_{j \in [q]}$ vecteurs propres D -orthonormés de WD associés aux q premières valeurs propres. C'est une base orthonormale de F^q, I_q
 $U(p \times q)$ matrice des vecteurs $(u_j)_{j \in [q]}$
 $\Gamma(n \times q)$ matrice des vecteurs $(\gamma_j)_{j \in [q]}$
 $C = \Gamma \Delta$ tableau ($n \times q$) des composantes principales du triplet (X, M, D)
 φ_{DM} produit scalaire dans l'espace des applications linéaires de E^* dans F

b) Pour une famille de triplets indexés

K ensemble des m situations donnant un triplet (X_k, M_k, D) ou (X_k, M, D_k)
 X_{IJK} cube de données obtenu si $J_k = J$ (ou $I_k = I$)

X_a, X_e	le tableau conjoint aligné $(X_1 \vdots \dots \vdots X_m)$, ou empilé $\begin{pmatrix} X_1 \\ \dots \\ \vdots \\ \dots \\ X_m \end{pmatrix}$
$k, k \in K$	indice ou exposant de toute notation relative au tableau unique signifiant que l'objet est pris en situation k
ρ_k	poinds des tableaux X_k
$E^p = (\mathbf{R}^p, \mathcal{M})$	espace somme directe orthogonale des $E_k : p = \sum p_k, \mathcal{M} = \text{diag}(M_k)$
\tilde{X}_k, \tilde{N}_k	tableau $(0 \vdots \dots \vdots X_k \vdots \dots \vdots 0)$, $(n \times p)$ qui définit \tilde{N}_k
\sim (de \tilde{N}_k, \dots)	signifie "qu'on plonge" (ici le nuage N_k) dans E^p
\wedge (de \tilde{X}_k, \dots)	signifie qu'on projette sur le sous espace de E^p spécifique à la méthode, excepté pour $\hat{C}_k = Y_k Q_E$ et \hat{C}, \hat{Y} qui sont des moyennes de \hat{C}_k ou \hat{Y}_k
$\bar{\cdot}$ (de \bar{Y}, \dots)	précise qu'il s'agit d'un tableau moyen
X_{ρ}, r	tableau moyen des \tilde{X}_k , rang $X_{\rho} = r$
$W_0, (C_0, \Lambda_0)$	le W du triplet $(X_{\rho}, \mathcal{M}, D)$, et les éléments propres de $W_0 D$
$C = \sqrt{m} C_0$	le tableau compromis $(n \times r)$
E_C	sous espace Compromis engendré par les individus moyens
$(u_j)_{j \in [r]}$	base orthonormale de E_C portant les axes principaux du nuage moyen
q	$q = \sup(q_k)$, (q_k est le rang de X_k)
$(e_j)_{j \in [q]}$	le Repère Equisecteur (définitions 2 et 3)
E_E	sous espace Equisecteur engendré par le Repère Equisecteur
$(v_j)_{j \in [q]}$	le Repère Déformant (définition 7)
E_D	sous espace Déformant engendré par le Repère Déformant
(Y_k, I, D)	triplet équivalent au triplet (X_k, M_k, D) , Y_k contient les coordonnées de N_k dans la base $(v_j)_{j \in [q]}$
A_k, a_{kj}	matrice diagonale s.d.p $(q \times q)$. $A_k = \text{diag}(a_{kj})$
Y, ε_k	matrices $(n \times q)$ telles que $\varepsilon_k = Y_k - Y A_k$ et $\sum \ \varepsilon_k\ ^2$ minimale
E_k^c, E^{p^c}	espace E_k éventuellement complété à la dimension q , et $\bigoplus E_k^c$
E_k^q, E^{mq}	sous espace de E_k^c de dimension q contenant E_k^{qk} , et $\bigoplus E_k^q$
C_k^c	le tableau C_k $(n \times q_k)$, complété par $q - q_k$ colonnes de 0
C_k^q	le tableau C_k $(n \times q_k)$, complété par $q - q_k$ colonnes de 0, l'ordre des colonnes permet de vérifier $C_k^q = C A_k$
P_k, P_k^c	le tableau $C_k^c D C \Lambda^{-1}$ $(q_k \times q)$, et $C_k^{c'} D C \Lambda^{-1}$ $(q \times q)$
Q, Q_k, Q_E	matrices orthogonales telles que : $C = Y Q, C_k = Y_k Q_k, \bar{C} = \bar{Y} Q_E$

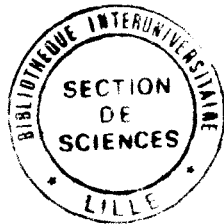
En annexe, on peut trouver des notations différentes, le sens en est alors précisé.

BIBLIOGRAPHIE.

- BOUROCHE J.M. (1975)** - Analyse des données ternaires : la double analyse en composantes principales. *Thèse de 3^{ième} cycle. Université de Paris VI.*
- BOUROCHE J.M. & SAPORTA G. (1980)** - L'Analyse des Données. *Que Sais-je ?, PUF - PARIS.*
- CAILLIEZ F. & PAGES J.P. (1976)** - Introduction à l'Analyse des Données. *SMASH - PARIS.*
- CARROLL J.D. (1968)** - A generalisation of canonical correlation analysis to three or more sets of variables. *Proc. 76th Conv. Amer. Psych. Assoc. p.227-228.*
- CARROLL J.D. & CHANG, J.J. (1970)** - Analysis of individual differences in multidimensional scaling via an n-way generalizations of "Eckart-Young" decomposition. *Psychometrika, 35, p. 283-319.*
- CASIN Ph. (1985)** - L'Analyse Canonique Généralisée et l'analyse des tableaux évolutifs. *Thèse de 3^{ième} cycle. Université L. Pasteur Strasbourg.*
- CASIN Ph. & TURLOT J.C. (1986)** - Une présentation de l'Analyse Canonique Généralisée dans l'espace des individus. *R.S.A. 1986 - Vol. XXXV n° 3, p. 65-75.*
- COPPI R. & BOLASCO S. (1989)** - Multiway data analysis. Volume originated from Multiway'88 international meeting on the multiway data matrices. CNR Rome Italie 28-30 Mars. *North Holland.*
- DE LEEUW J. & PRUZANSKY S. (1978)** - A new computational method to fit the weighted euclidean distance model. *Psychometrika, 43, n° 4, p. 479-490.*
- DE LEEUW J. & HEISER W. (1982)** - Theory of Multidimensionnal Scaling. in P.R. Krishnaiah & L.N. Kanal (Eds.) : *Handbook of Statistics, Vol. 2. North Holland, p. 285-316.*
- ECKART C. & YOUNG G. (1936)** - The approximation of one matrix by another of low rank. *Psychometrika, 1, p. 211-218.*
- ESCOFIER B. & PAGES J. (1983)** - L'Analyse Factorielle Multiple. Une méthode de comparaison de groupes de variables. in Diday E. et al. (Eds) : *Data Analysis and Informatics III. North Holland 1984.*
- ESCOFIER B. & PAGES J. (1988)** - Analyses Factorielles Simples et Multiples. *DUNOD - PARIS.*

- ESCOUFIER Y. et al. (1985)** - Numéro spécial sur les tableaux à 3 dimensions et données évolutives. *Statistique et Analyse des Données. Juin 1985. (SAD Vol. 10, n° 1).*
- GLACON F. (1981)** - Analyse conjointe de plusieurs matrices de données : comparaison de méthodes. *Thèse de 3^{ème} cycle - Université Médicale de Grenoble.*
- HARSHMAN R.A. (1970)** - Foundations of the PARAFAC procedure : models and conditions for an "explanatory" multi-mode analysis. *UCLA Working Papers in Phonetics, 16, p. 1-84.*
- HARSHMAN R.A. & LUNDY, M.E. (1984)** - The PARAFAC model for Three-Way Factor Analysis and Multidimensional Scaling. in H.G. Law, C.W. Snyder, J.A. Hattie, and R.P. McDonald (Eds.) : *Research methods for multimode data analysis (p. 122-215). New York : Praeger.*
- JOLLIFFE J.T. (1986)** - Principal Component Analysis. *New York : Springer-Verlag.*
- KIERS H.A.L. (1987)** - Comparaison de "anglo-saxon" and "French" three way methods and their application to the analysis of nominal variables. *présenté à la workshop "données évolutives" Data Analysis and Informatics V.*
- KIERS H.A.L. (1988)** - Hierarchical relations between three-way methods. *présenté aux "XX Journées de Statistique" Grenoble 2-6-1988.*
- KIERS H.A.L. (1989)** - Three ways methods for the analysis of qualitative and quantitative two ways data. *Leiden : DSWO-press.*
- KRISTOF W. (1970)** - A theorem on the trace of certain matrix products and some applications. *Journal of Mathematical psychology, Vol. 7, p.515-530.*
- KROONENBERG P.M. (1983)** - Three mode principal component analysis : Theory and applications. *Leiden : DSWO-press.*
- LAFAYE J.Y. (1985)** - Application du modèle INDSCAL. in Escoufier Y. et al. : *Vol. 10, n° 1 p. 83-102.*
- LAFOSSE R. (1985)** - Analyses procustéennes de deux tableaux. *Thèse de 3^{ème} cycle. Université P. Sabatier Toulouse.*
- LAVIT C. (1985)** - Application de la méthode Statis. in Escoufier Y. et al. : *Vol. 10, n° 1 p. 103-116.*
- LAVIT C. (1988)** - Analyse conjointe de tableaux quantitatifs. Méthode et programmes. *MASSON - PARIS.*
- LE CALVE G. (1976)** - Problèmes d'Analyse des données. *Thèse d'Etat. Université de Rennes I.*
- L'HERMIER des plantes H. (1976)** - S.T.A.T.I.S. : théorie et application d'une méthode d'analyse conjointe. *Thèse de 3^{ème} cycle. Université des Sciences et Techniques du Languedoc.*

- MOUTTET F. (1982)** - Comparaison de deux tableaux par la méthode procruste. *Thèse de 3^{ème} cycle. Université de Paris VI.*
- POLIT M.O. E. (1986)** - Une n-ACP d'un hypercube de données. *Thèse de 3^{ème} cycle - Université de Grenoble II.*
- SAPORTA G. (1975)** - Liaison entre plusieurs ensembles de variables et codage de données qualitatives. *Thèse de 3^{ème} cycle. Université de Paris VI.*
- SAPORTA G. (1990)** - Probabilité, Analyse des données et Statistique. *TECHNIP - PARIS .*
- TEN BERGE J.M.F. (1977)** - Orthogonal Procrustes rotation for two or more matrices. *Psychometrika, 42, p. 267-276.*
- TEN BERGE J.M.F. (1983)** - A generalization of Kristof's theorem on the trace of certain matrix products. *Psychometrika, 48, p. 519-523.*
- TEN BERGE J.M.F. (1986)** - A general solution for the Maxbet problem. in J. de Leeuw, W.J. Heiser, J. Meulman & F. Critchley (Eds.) : *Multidimensional Data Analysis. Leiden : DSWO-press, p. 81-87.*
- TEN BERGE J.M.F. (1988)** - Generalized approaches to the Maxbet problem and the maxdiff problem, with applications to canonical correlations. *Psychometrika, 53, n° 4, p. 487-494.*
- TEN BERGE J.M.F. & KNOL D.L. (1984)** - Orthogonal rotations to maximal agreement for two or more matrices of different column orders. *Psychometrika, 49, p. 49-55.*
- TUCKER L.R. (1966)** - Some mathematical notes on three-mode factor analysis. *Psychometrika, 31 n° 3, p. 279-311.*
- TUCKER L.R. (1972)** - Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika, 37 n° 1, p. 3-27.*



Résumé :

Pour une famille de tableaux contenant les observations de différents paquets de variables quantitatives sur un ensemble d'individus, nous posons le problème du positionnement simultané des nuages d'individus, appelé problème de l'intrastructure.

Nous définissons l'Analyse Factorielle conJointe (AFJ) qui, contrairement aux méthodes existantes, généralise l'Analyse en Composantes Principales (ACP) d'un tableau au cas d'une famille de tableaux. En effet, l'intrastructure contient une image euclidienne équivalente des nuages initiaux, et le positionnement simultané respecte les règles de la dualité.

Deux approches complémentaires sont proposées. Une définition géométrique permet de présenter plusieurs méthodes dans un espace commun où elles apparaissent comme méthodes de projection sur tel sous espace, et ainsi d'étudier leurs relations. Une définition analytique conduit à un algorithme qui prouve l'existence d'une solution et la détermine.

Sur trois exemples nous comparons les résultats des différentes méthodes. Nous précisons également jusqu'à quel point ces méthodes sont utilisables pour positionner les individus, mais aussi pour analyser les relations entre les paquets de variables ou entre les situations.

Mots clés :

Compromis - Interstructure - Intrastructure - Indscal - Analyse Factorielle Multiple - Analyse Canonique - Analyse Procuste - Composantes conJointes.