Analyse conjointe de plusieurs matrices de données : comparaison de différentes méthodes

Frédérique Glacon

THESE

présentée à

l' Université Scientifique et Médicale de Grenoble

pour obtenir le grade de
DOCTEUR DE 3ème CYCLE
spécialité : Statistiques

par

Frédérique GLACON

ANALYSE CONJOINTE DE PLUSIEURS MATRICES DE DONNEES

COMPARAISON DE DIFFERENTES METHODES

Thèse soutenue le 17 juin 1981 devant la Commission d'Examen :

Monsieur	J.R. BARRA	Président
Messieurs	D. DROUET D'AUBIGNY	Examinateurs
	Y. ESCOUFIER	
	G. ROMIER	
	B. VAN CUTSEM	

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Monsieur Gabriel CAU : Président
Monsieur Joseph KLEIN : Vice-Président

MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

PROFESSEURS TITULAIRES

MM.	AMBLARD Pierre	Clinique de dermatologie
	ARNAUD Paul	Chimie
	ARVIEU Robert	I.S.N.
	AUBERT Guy	Physique
	AYANT Yves	Physique approfondie
Mme	BARBIER Marie-Jeanne	Electrochimie
MM.	BARBIER Jean-Claude	Physique expérimentale
	BARBIER Reynold	Géologie appliquée
	BARJON Robert	Physique nucléaire
	BARNOUD Fernand	Biosynthèse de la cellulose
	BARRA Jean-René	Statistiques
	BARRIE Joseph	Clinique chirurgicale A
	BEAUDOING André	Clinique de pédiatrie et puériculture
	BELORIZKY Elie	Physique
	BARNARD Alain	Mathématiques pures
Mme	BERTRANDIAS Françoise	Mathématiques pures
MM.	BERTRANDIAS Jean-Paul	Mathématiques pures
	BEZES Henri	Clinique chirurgicale et traumatologie
	BLAMBERT Maurice	Mathématiques pures
	BOLLIET Louis	Informatique (I.U.T. B)
	BONNET Jean-Louis	Clinique ophtalmologie
	BONNET-EYMARD Joseph	Clinique hépato-gastro-entérologie
Mme	BONNIER Marie-Jeanne	Chimie générale
MM.	BOUCHERLE André	Chimie et toxicologie
	BOUCHEZ Robert	Physique nucléaire
	BOUSSARD Jean-Claude	Mathématiques appliquées
	BOUTET DE MONVEL Louis	Mathématiques pures
	BRAVARD Yves	Géographie
	CABANEL Guy	Clinique rhumatologique et hydrologique
	CALAS François	Anatomie
	CARLIER Georges	Biologie végétale
	CARRAZ Gilbert	Biologie animale et pharmacodynamie

.../...

MM.	CAU Gabriel	Médecine légale et toxicologie
	CAUQUIS Georges	Chimie organique
	CHABAUTY Claude	Mathématiques pures
	CHARACHON Robert	Clinique ot-rhino-laryngologique
	CHATEAU Robert	Clinique de neurologie
	CHIBON Pierre	Biologie animale
	COEUR André	Pharmacie chimique et chimie analytique
	COUDERC Pierre	Anatomie pathologique
	DEBELMAS Jacques	Géologie générale
	DEGRANGE Charles	Zoologie
	DELORMAS Pierre	Pneumophtisiologie
	DEPORTES Charles	Chimie minérale
	DESRE Pierre	Métallurgie
	DODU Jacques	Mécanique appliquée (I.U.T. I)
	DOLIQUE Jean-Michel	Physique des plasmas
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	FONTAINE Jean-Marc	Mathématiques pures
	GAGNAIRE Didier	Chimie physique
	GALVANI Octave	Mathématiques pures
	GASTINEL Noël	Analyse numérique
	GAVEND Michel	Pharmacologie
	GEINDRE Michel	Electroradiologie
	GERBER Robert	Mathématiques pures
	GERMAIN Jean-Pierre	Mécanique
	GIRAUD Pierre	Géologie
	JANIN Bernard	Géographie
	KAHANE André	Physique générale
	KLEIN Joseph	Mathématiques pures
	KOSZUL Jean-Louis	Mathématiques pures
	KRAVTCHENKO Julien	Mécanique
	LACAZE Albert	Thermodynamique
	LACHARME Jean	Biologie végétale
Mme	LAJZEROWICZ Janine	Physique
MM.	LAJZEROWICZ Joseph	Physique
	LATREILLE René	Chirurgie générale
	LATURAZE Jean	Biochimie pharmaceutique
	LAURENT Pierre	Mathématiques appliquées
	LEDRU Jean	Clinique médicale B
	LE ROY Philippe	Mécanique (.I.U.T. I)

.../...

MM.	LLIBOUTRY Louis	Géophysique
	LOISEAUX Jean-Marie	Sciences nucléaires
	LONGEQUEUE Jean-Pierre	Physique nucléaire
	LOUP Jean	Géographie
Mlle	LUTZ Elisabeth	Mathématiques pures
MM.	MALINAS Yves	Clinique obstétricale
	MARTIN-NOEL Pierre	Clinique cardiologique
	MAYNARD Roger	Physique du solide
	MAZARE Yves	Clinique Médicale A
	MICHEL Robert	Minéralogie et pétrographie
	MICOUD Max	Clinique maladies infectieuses
	MOURIQUAND Claude	Histologie
	MOUSSA André	Chimie nucléaire
	NEGRE Robert	Mécanique
	NOZIERES Philippe	Spectrométrie physique
	OZENDA Paul	Botanique
	PAYAN Jean-Jacques	Mathématiques pures
	PEBAY-PEYROULA Jean-Claude	Physique
	PERRET Jean	Séméiologie médicale (neurologie)
	RASSAT André	Chimie systématique
	RENARD Michel	Thermodynamique
	REVOL Michel	Urologie
	RINALDI Renaud	Physique
	DE ROUGEMONT Jacques	Neuro-Chirurgie
	SARRAZIN Roger	Clinique chirurgicale B
	SEIGNEURIN Raymond	Microbiologie et hygiène
	SENGEL Philippe	Zoologie
	SIBILLE Robert	Construction mécanique (I.U.T. I)
	SOUTIF Michel	Physique générale
	TANCHE Maurice	Physiologie
	VAILLANT François	Zoologie
	VALENTIN Jacques	Physique nucléaire
Mme	VERAIN Alice	Pharmacie galénique
MM.	VERAIN André	Physique biophysique
	VEYRET Paul	Géographie
	VIGNAIS Pierre	Biochimie médicale

.../...

PROFESSEURS ASSOCIES

MM.	CRABBE Pierre	CERMO
	SUNIER Jules	Physique

PROFESSEURS SANS CHAIRE

Mlle	AGNIUS-DELORS Claudine	Physique pharmaceutique
	ALARY Josette	Chimie analytique
MM.	AMBROISE-THOMAS Pierre	Parasitologie
	ARMAND Gilbert	Géoφ φaphie
	BENZAKEN Claude	Mathématiques appliquées
	BIAREZ Jean-Pierre	Mécanique
	BILLET Jean	Géographie
	BOUCHET Yves	Anatomie
	BRUGEL Lucien	Energétique (I.U.T. I)
	BUISSON René	Physique (I.U.T. I)
	BUTEL Jean	Orthopédie
	COHEN-ADDAD Jean-Pierre	Spectrométrie physique
	COLOMB Maurice	Biochimie médicale
	CONTE René	Physique (I.U.T. I)
	DELOBEL Claude	M.I.A.G.
	DEPASSEL Roger	Mécanique des fluides
	GAUTRON René	Chimie
	GIDON Paul	Géologie et minéralogie
	GLENAT René	Chimie organique
	GROULADE Joseph	Biochimie médicale
	HACQUES Gérard	Calcul numérique
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Hygiène et médecine préventive
	IDELMAN Simon	Physiologie animale
	JOLY Jean-René	Mathématiques pures
	JULLIEN Pierre	Mathématiques appliquées
Mme	KAHANE Josette	Physique
MM.	KRAKOWIACK Sacha	Mathématiques appliquées
	KUHN Gérard	Physique (I.U.T. I)
	LUU DUC Cuong	Chimie organique - pharmacie
	MICHOULIER Jean	Physique (I.U.T. I)
Mme	MINIER Colette	Physique (I.U.T. I)

.../...

MM.	PELMONT Jean	Biochimie
	PERRIAUX Jean-Jacques	Géologie et minéralogie
	PFISTER Jean-Claude	Physique du solide
Mlle	PIERY Yvette	Physiologie animale
MM.	RAYNAUD Hervé	M.I.A.G.
	REBECQ Jacques	Biologie (CUS)
	REYMOND Jean-Charles	Chirurgie générale
	RICHARD Lucien	Biologie végétale
Mme	RINAUDO Marguerite	Chimie macromoléculaire
MM.	SARROT-REYNAULD Jean	Géologie
	SIROT Louis	Chirurgie générale
Mme	SOUTIF Jeanne	Physique générale
MM.	STIEGLITZ Paul	Anesthésiologie
	VIALON Pierre	Géologie
	VAN CUTSEM Bernard	Mathématiques appliquées

MAITRES DE CONFERENCES ET MAITRES DE CONFERENCES AGREGES

MM.	ARMAND Yves	Chimie (I.U.T. I)
	BACHELOT Yvan	Endocrinologie
	BARGE Michel	Neuro-chirurgie
	BEGUIN Claude	Chimie organique
Mme	BERIEL Hélène	Pharmacodynamie
MM.	BOST Michel	Pédiatrie
	BOUCHARLAT Jacques	Psychiatrie adultes
Mme	BOUCHE Liane	Mathématiques (CUS)
MM.	BRODEAU François	Mathématiques (I.U.T. B) (Personne étrangère habilitée à être directeur de thèse)
	BERNARD Pierre	Gynécologie
	CHAMBAZ Edmond	Biochimie médicale
	CHAMPETIER Jean	Anatomie et organogénèse
	CHARDON Michel	Géographie
	CHERADAME Hervé	Chimie papetière
	CHIAVERINA Jean	Biologie appliquée (EFP)
	COLIN DE VERDIERE Yves	Mathématiques pures
	CONTAMIN Charles	Chirurgie thoracique et cardio-vasculaire
	CORDONNER Daniel	Néphrologie
	COULOMB Max	Radiologie
	CROUZET Guy	Radiologie

.../...

MM.	CYROT Michel	Physique du solide
	DENIS Bernard	Cardiologie
	DOUCE Roland	Physiologie végétale
	DUSSAUD René	Mathématiques (CUS)
Mme	ETERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine légale
	FAURE Gilbert	Urologie
	GAUTIER Robert	Chirurgie générale
	GIDON Maurice	Géologie
	GROS Yves	Physique (I.U.T. I)
	GUIGNIER Michel	Thérapeutique
	GUITTON Jacques	Chimie
	HICTER Pierre	Chimie
	JALBERT Pierre	Histologie
	JUNIEN-LAVILLAVROY Claude	O.R.L.
	KOLODIE Lucien	Hématologie
	LE NOC Pierre	Bactériologie-virologie
	MACHE Régis	Physiologie végétale
	MAGNIN Robert	Hygiène et médecine préventive
	MALLION Jean-Michel	Médecine du travail
	MARECHAL Jean	Mécanique (I.U.T. I)
	MARTIN-BOUYER Michel	Chimie (CUS)
	MASSOT Christian	Médecine interne
	NEMOZ Alain	Thermodynamique
	NOUGARET Marcel	Automatique (I.U.T. I)
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (I.U.T. B) (Personnalité étrangère habilitée à être directeur de thèse)
	PEFFEN René	Métallurgie (I.U.T. I)
	PERRIER Guy	Géophysique-glaciologie
	PHELIP Xavier	Rhumatologie
	RACHALL Michel	Médecine interne
	RACINET Claude	Gynécologie et obstétrique
	RAMBAUD Pierre	Pédiatrie
	RAPHAEL Bernard	Stomatologie
Mme	RENAUDET Jacqueline	Bactériologie (pharmacie)
MM.	ROBERT Jean-Bernard	Chimie-physique
	ROMIER Guy	Mathématiques (I.U.T. B) (Personnalité étrangère habilitée à être directeur de thèse)
	SAKAROVITCH Michel	Mathématiques appliquées

.../...

MM.	SCHAERER René	Cancérologie
Mme	SEIGLE-MURANDI Françoise	Crytogamie
MM.	STOEBNER Pierre	Anatomie pathologie
	STUTZ Pierre	Mécanique
	VROUSOS Constantin	Radiologie

MAITRES DE CONFERENCES ASSOCIES

MM.	DEVINE Roderick	Spectro Physique
	KANEKO Akira	Mathématiques pures
	JOHNSON Thomas	Mathématiques appliquées
	RAY Tuhina	Physique

MAITRE DE CONFERENCES DELEGUE

M.	ROCHAT Jacques	Hygiène et hydrologie (pharmacie)

Fait à Saint Martin d'Hères, novembre 1977

Je tiens à remercier Monsieur le Professeur J.R. BARRA qui me fait le grand honneur de présider le jury de cette thèse, ainsi que Messieurs B. VAN CUTSEM, G.ROMIER, et D. DROUET D'AUBIGNY qui ont aimablement accepté d'y participer .

J'exprime ma vive reconnaissance à Monsieur le Professeur Y. ESCOUFIER qui m'a proposé le sujet de cette thèse et m'a patiemment guidée tout au long de ce travail .

Je voudrais aussi associer dans un même remerciement amical tous mes camarades du C.R.I.G. à Montpellier ,ainsi que Madame LOUCHE qui m'a aidée à la frappe de cette thèse .

SOMMAIRE

INTRODUCTION GENERALE :

De nombreuses méthodes ont été proposées ces dernières années pour l'étude conjointe de plusieurs matrices de données, quantitatives ou qualitatives, afin de généraliser à plusieurs tableaux les principales techniques de l'Analyse des Données.

Le but de ce travail n'est pas d'ajouter une nouvelle méthode à celles déjà existantes, mais de les présenter de manière homogène, afin de pouvoir les comparer. Une comparaison est faite, en particulier, entre les approches "Française"et "Américaine" du problème, qui sont assez différentes.

- Dans un premier chapitre, on expose quelques résultats bien connus en Analyse des Données, qui seront utilisés tout au long de ce travail, et on met en place les notations employées par la suite. Les résultats sont présentés sous la forme la plus générale possible (introduction de métriques quelconques sur tous les espaces considérés, même si habituellement on n'utilise que la métrique identité).
Les démonstrations ne sont données que pour les théorèmes plus généraux que ceux habituellement utilisés.

- Dans le deuxième chapitre, on montre à quels types de situations on peut être confronté, et comment caractériser ces situations.

- Le troisième chapitre est consacré à un exposé synthétique des méthodes le plus couramment employées en France pour l'étude conjointe de plusieurs matrices de données. On dégage la structure commune de ces méthodes, et on compare leurs différentes propriétés, ce qu'elles mettent en évidence. Des justifications nouvelles seront ainsi apportées à certaines pratiques. Les parties les plus nouvelles concernent la comparaison des"compromis" enIII-3, ainsi que l'exposé des propriétés des différentes "intrastructures" en III-4-1, où une nouvelle méthode ayant d'autres propriétés optimales est suggérée.

- Les chapitres IV et V seront consacrés à la présentation des méthodes utilisées essentiellement aux Etats-Unis et en Grande-Bretagne pour étudier les mêmes situations : on met en évidence les propriétés mathématiques de ces méthodes, et on les compare aux approches précédentes.

- Au chapitre IV, après avoir exposé l'utilisation classique des modèles INDSCAL et IDIOSCAL, on montre dans quelles conditions ces modèles sont applicables et on propose (en 3-2) un nouveau moyen, inspiré des méthodes précédentes, pour tester leur validité, après avoir montré l'inexactitude d'une autre pratique.

- Au chapitre V, on résout le problème procustéen orthogonal dans un cadre plus général que celui existant (tableaux n'ayant pas le même nombre de variables), et on montre que la solution rejoint des pratiques habituelles et permet de les justifier.

L'Analyse Procustéenne Généralisée est présentée de façon à pouvoir établir immédiatement le parallèle entre les méthodes Procustéennes et l'approche "Interstructure-Compromis-Intrastructure" décrite au chapitre III. En outre, l'Analyse Procustéenne permet d'apporter une justification supplémentaire à la méthode "STATIS", montrée en 3-3.

*

* *

NOTATIONS :

Les notations utilisées seront celles, assez généralement employées, du livre de F.CAILLIEZ et J.P.PAGES (4) .

Le triplet (X,M,D) :

Soit X un tableau de données résultant de la mesure de p variables réelles sur n individus (tableau à p lignes et n colonnes) .

Soit $E = \mathbb{R}^p$, muni de la base canonique $\{e_1, e_2, \ldots\ldots, e_p\}$.
Soit $F = \mathbb{R}^n$, muni de la base canonique $\{f_1, f_2, \ldots\ldots, f_n\}$.
Le dual de E , E^* , a pour base canonique $\{e_1^*, e_2^*, \ldots, e_p^*\}$.
Le dual de F , F^* , a pour base canonique $\{f_1^*, f_2^*, \ldots\ldots, f_n^*\}$.

Le tableau X est assimilé à la matrice de l'application linéaire:

$$\mathbb{R}^{n^*} \longrightarrow \mathbb{R}^p$$

$$f_j^* \longmapsto \begin{bmatrix} x_{1j} \\ x_{2j} \\ \cdot \\ x_{pj} \end{bmatrix} \qquad (j^{\text{ème}} \text{ colonne du tableau X})$$

On notera aussi X cette application linéaire.

De façon plus générale, on notera de manière identique, lorsqu'il n'y a pas de confusion possible, une application linéaire d'un espace vectoriel réel dans un autre et sa matrice dans les bases canoniques.

Une métrique euclidienne sur un espace vectoriel réel E est définie par :

- la donnée d'une forme bilinéaire symétrique définie positive f; on note alors M la matrice pxp des $f(e_i, e_j)$
- ou par la donnée de l'application linéaire g :

$$E \longrightarrow E^*$$
$$x \longmapsto (y \longmapsto f(x,y))$$

M est aussi la matrice de g dans les bases canoniques.

Lorsqu'il n'y a pas de confusion possible, on notera indifféremment M :

- la forme f
- l'isomorphisme g
- la matrice symétrique définie positive M .

On notera aussi $\langle x,y\rangle_M = f(x,y)$

$\|x\|^2_M = f(x,x)$ pour tout x et tout y de E.

La donnée du tableau X va de pair avec celle de deux métriques euclidiennes :

- l'une, de matrice M, p x p, permet de mesurer les proximités entre individus, considérés comme éléments de $E = \mathbb{R}^p$.
- l'autre, de matrice N, n x n, permet de mesurer les proximités entre variables, considérées comme éléments de $F = \mathbb{R}^n$.

Généralement on a N = D, matrice diagonale définie positive, dont les éléments diagonaux sont les poids affectés aux n individus.

X est centré pour D si $XD\underset{\sim}{1} = 0$, où $\underset{\sim}{1} = \begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$.

Le schéma de dualité :

On a le schéma suivant :

$$\begin{array}{ccc} E = \mathbb{R}^p & \xleftarrow{\quad X \quad} & F^* \\ M \downarrow\uparrow V & & W \downarrow\uparrow D \\ E^* & \xrightarrow[\quad X' \quad]{} & F = \mathbb{R}^n \end{array}$$

X' est la transposée de X ; $X'(e^*_i)_j = x_{ij}$.

Les variables peuvent être représentées :

- dans F par les vecteurs $x^i = \begin{bmatrix}x_{i1}\\\vdots\\x_{in}\end{bmatrix}$,valeurs prises par la variable pour les n individus .

- dans E par les vecteurs e_i
- dans E^* par les vecteurs e_i^*.

Les individus peuvent être représentés :

- dans E par les vecteurs $x_j = \begin{bmatrix} x_{1j} \\ \vdots \\ x_{pj} \end{bmatrix}$, valeurs prises par les p variables sur l'individu j .
- dans F par les vecteurs f_j
- dans F^* par les vecteurs f_j^* .

On définit V et W par :

$$V = XDX' \qquad \text{et} \qquad W = X'MX$$

W est la matrice des produits scalaires entre individus de E .
V est la matrice des produits scalaires entre variables de F .

Si X est centré, V est la matrice des variances et covariances entre les p variables du tableau X .

I : RESULTATS PRÉLIMINAIRES :

1 / THEOREME DE DECOMPOSITION :

Soit E_1 un espace vectoriel euclidien réel, de dimension p_1, de métrique M_1 , et soit E_2 un espace vectoriel euclidien réel, de dimension p_2, de métrique M_2. Soit X la matrice d'une application linéaire de E_2^* dans E_1, de rang r .

$X' M_1 X M_2$ est M_2-symétrique,positive;ses valeurs propres sont positives; on les note $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_{p_2} = 0$ .Soit $\{V_1, V_2, \ldots, V_{p_2}\}$ un système M_2-orthonormal de vecteurs propres associés. On pose :

$$V = (V_1 \quad V_2 \ldots\ldots \quad V_r)$$

$$\Delta = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \sqrt{\lambda_2} & \\ & & \ddots \\ 0 & & \sqrt{\lambda_r} \end{pmatrix}$$

$$U = X M_2 V \Delta^{-1}$$

$$\begin{array}{ccc} E_1 & \xleftarrow{\quad X \quad} & E_2^* \\ M_1 \downarrow & & \uparrow M_2 \\ E_1^* & \xrightarrow[\quad X' \quad]{} & E_2 \end{array}$$

Théorème 1 : $X = U \Delta V'$,avec $U' M_1 U = I_r$, $V' M_2 V = I_r$,Δ est une matrice diagonale positive.

Démonstration : Soit $\tilde{V} = (V_1 \ldots . V_r \ldots\ldots V_{p_2})$, $\tilde{V}'$ est une matrice carrée telle que $\tilde{V}' M_2 \tilde{V} = I_{p_2}$, $M_2 \tilde{V}$ est l'inverse de $\tilde{V}'$ donc $M_2 \tilde{V} \tilde{V}' = I_{p_2}$.

D'autre part, pour tout $i > r$, $X' M_1 X M_2 V_i = 0$, donc

$\| X M_2 V_i \|^2_{M_1} = 0$,et $X M_2 V_i = 0$.

Donc $X = X I_{p_2} = X M_2 \tilde{V} \tilde{V}' = X M_2 \sum_{i=1}^{p_2} V_i V_i' = X M_2 \sum_{i=1}^{r} V_i V_i' = X M_2 V V' = U \Delta V'$

et $U' M_1 U = \Delta^{-1} V' M_2 X' M_1 X M_2 V \Delta^{-1} = \Delta^{-1} V' M_2 V \Delta^2 \Delta^{-1} = I_r$. .

Conséquence : Les colonnes de U : $U_1, U_2, \ldots, U_r$, sont un système M_1-orthonormal de vecteurs propres de $X M_2 X' M_1$, associés à $\lambda_1, \lambda_2, \ldots, \lambda_r$, et $V = X' M_1 U \Delta^{-1}$.

En effet :

$$X M_2 X' M_1 U_i = X M_2 X' M_1 \left(\frac{1}{\sqrt{\lambda_i}} X M_2 V_i \right) = \frac{1}{\sqrt{\lambda_i}} X M_2 (\lambda_i V_i) = \lambda_i \left(\frac{1}{\sqrt{\lambda_i}} X M_2 V_i \right) = \lambda_i U_i$$

et

$$X' M_1 U \Delta^{-1} = V \Delta U' M_1 U \Delta^{-1} = V .$$

2 / PRODUIT SCALAIRE DE MATRICES OU D'APPLICATIONS LINEAIRES :

Soit (E_1, M_1) et (E_2, M_2) deux espaces vectoriels euclidiens définis comme précédemment. Soit $\mathcal{M}_{p_1 \times p_2}$ l'espace vectoriel des matrices $p_1 \times p_2$ isomorphe à l'espace vectoriel des applications linéaires de $R^{p_2 *}$ dans R^{p_1} .

Proposition : L'application $\varphi_{M_1, M_2} : \mathcal{M}_{p_1 \times p_2} \times \mathcal{M}_{p_1 \times p_2} \longrightarrow R$

$$(A, B) \longmapsto \mathrm{Tr}\,(A' M_1 B M_2)$$

est bilinéaire, symétrique, définie positive .
(Tr désigne la trace de la matrice) .

Démonstration : La bilinéarité et la symétrie résultent des propriétés de la trace.

Soit $A \in \mathcal{M}_{p_1 \times p_2}$; $\varphi_{M_1, M_2}(A, A) = \mathrm{Tr}(A' M_1 A M_2)$

Si $M_1 = L_1 L_1'$, où L_1 est une matrice $p_1 \times p_1$ de rang p_1, et $M_2 = L_2 L_2'$, où L_2 est une matrice $p_2 \times p_2$ de rang p_2, on a :

$$\mathrm{Tr}(A' M_1 A M_2) = \mathrm{Tr}(A' L_1 L_1' A L_2 L_2') = \mathrm{Tr}(L_1' A L_2 \, L_2' A' L_1) = \sum_{i=1}^{p_1} \sum_{j=1}^{p_2} (L_1' A L_2)_{ij}^2 \geq 0$$

et

$$Tr(A'M_1AM_2) = 0 \iff L_1'AL_2 = 0 \iff A = 0$$

$\varphi_{M_1M_2}$ est définie positive .

Donc pour tout choix de métriques M_1 sur E_1, et M_2 sur E_2, on peut définir un produit scalaire $\varphi_{M_1M_2}$, ce qui permettra de définir la norme d'une application linéaire ou d'une matrice, ainsi qu'une distance euclidienne. Nous verrons au chapitre III comment utiliser ce produit scalaire, selon les espaces et les métriques considérés.

Remarque : $\mathcal{M}_{p_1xp_2}$ est isomorphe à $\mathbb{R}^{p_1xp_2}$, par l'isomorphisme qui, à toute matrice A, associe le vecteur $\tilde{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_{p_2} \end{pmatrix}$ obtenu en juxtaposant les colonnes de A.

$\varphi_{M_1M_2}$ peut alors être considérée comme une métrique sur $\mathbb{R}^{p_1xp_2}$, de matrice :

$$M_2 \otimes M_1 = \begin{bmatrix} (M_2)_{11}M_1 & (M_2)_{12}M_1 & \cdots\cdots\cdots & (M_2)_{1p_2}M_1 \\ (M_2)_{21}M_1 & (M_2)_{22}M_1 & \cdots\cdots\cdots & (M_2)_{2p_2}M_1 \\ \cdot & \cdot & \cdots\cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots\cdots & \cdot \\ (M_2)_{p_21}M_1 & (M_2)_{p_22}M_1 & \cdots\cdots\cdots & (M_2)_{p_2p_2}M_1 \end{bmatrix}$$

(Produit de Kronecker des matrices M_2 et M_1) .

3 / THEOREMES D'APPROXIMATIONS :

Les théorèmes ci-dessous seront donnés sans démonstration . On peut les trouver dans (24) ou (9) .

Théorème 2 : Soit (E,M) un espace vectoriel euclidien réel, de dimension p. Soit A une application linéaire de E^* dans E, symétrique, semi-définie positive. Soit $(e_1, e_2, \ldots\ldots, e_p)$ une base M-orthonormale de vecteurs propres de AM , associés aux valeurs propres $\lambda_1 \geq \lambda_2 \geq \ldots\ldots \geq \lambda_p \geq 0$.

On a $A = \sum_{i=1}^{p} \lambda_i e_i e_i'$. Alors , pour tout $r \leq p$,parmi les applications linéaires de E^* dans E , symétriques , semi-définies positives , de rang inférieur ou égal à r , le minimum de $\|A - B\|^2_{\varphi_{MM}}$ est atteint pour $B = \sum_{i=1}^{r} \lambda_i e_i e_i'$, ce minimum vaut $\sum_{i=r+1}^{p} \lambda_i^2$.

Théorème 3 : B, définie comme précédemment, réalise le maximum de $\frac{<A,B>_{\varphi_{MM}}}{\|B\|_{\varphi_{MM}}}$; ce maximum vaut $\sqrt{\lambda_1^2 + \ldots + \lambda_r^2}$.

Théorème 4 : Soit A une application linéaire de E_2^* dans E_1, où (E_1, M_1) et (E_2, M_2) sont deux espaces vectoriels euclidiens. Soit r le rang de A . On a selon le théorème 1

$$A = U\Delta V' = \sum_{i=1}^{r} \sqrt{\lambda_i}\, U_i V_i' .$$

Alors l'application linéaire B de E_2^* dans E_1 ,définie par $B = \sum_{i=1}^{s} \sqrt{\lambda_i}\, U_i V_i'$ réalise le minimum de $\|A - B\|^2_{\varphi_{M_1 M_2}}$, parmi les applications linéaires de rang $\leq s < r$. Ce minimum vaut $\sum_{i=s+1}^{r} \lambda_i$.

Une démonstration très complète de ce théorème est donnée dans (9).

Théorème 5 : Soit $X = U\Delta V'$, avec les notations du théorème 1.
Alors U_1 est le vecteur normé de E_1 réalisant le maximum de $\|X'M_1u\|^2_{M_2}$, et ce maximum vaut λ_1 .
Pour $i \leq r$, U_i est le vecteur normé de E_1, M_1-orthogonal à $U_1, U_2, \ldots$ et U_{i-1} , réalisant le maximum de $\|X'M_1u\|^2_{M_2}$ et ce maximum vaut λ_i .

Démonstration : $\|X'M_1u\|^2_{M_2} = u'(M_1XM_2X'M_1)u$ est maximum, avec la contrainte $u'M_1u = 1$, pour $M_1XM_2X'M_1u = kM_1u$, ou $XM_2X'M_1u = ku$, avec k maximum;

donc $u = U_1$, vecteur propre normé de $XM_2X'M_1$ associé à la plus grande valeur propre. De même, $|| X'M_1u||^2_{M_2}$ est maximum, avec les contraintes: $u'M_1u = 1$, $u'M_1U_1 = u'M_1U_2 = \ldots\ldots = u'M_1U_{i-1} = 0$,pour $u = U_i$.

$$|| X'M_1u ||^2_{M_2} = || \sqrt{\lambda_i} V_i ||^2_{M_2} = \lambda_i .$$

Cas particulier : Si M_2 est diagonale , d'éléments diagonaux $q_1, q_2, .., q_{p_2}$,

$$|| X'M_1u ||^2_{M_2} = \sum_{i=1}^{p_2} q_i \, \langle x_i, u \rangle^2_{M_1}$$,où x_i est la $i^{ème}$ colonne de X. Les U_i réalisent donc dans ce cas le maximum de :

$$\sum_{i=1}^{p_2} q_i \, \langle x_i, u \rangle^2_{M_1} .$$

Remarque : Etant donnée la symétrie du problème, les V_i réalisent de même le maximum de $|| XM_2v ||^2_{M_1}$, sous contrainte d'orthonormalité des solutions v .

De même, si M_1 est diagonale, d'éléments diagonaux $r_1, .., r_{p_1}$, les V_i réalisent le maximum de :

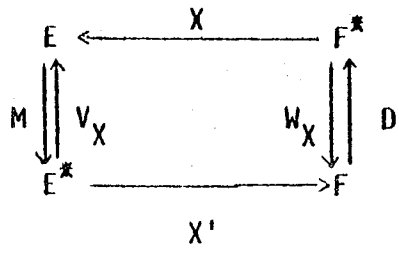
$$\sum_{j=1}^{p_1} r_j \, \langle x^j, v \rangle^2_{M_2}$$,où x^j est la $j^{ème}$ ligne de X .

4 / ANALYSE EN COMPOSANTES PRINCIPALES :

Nous allons appliquer les théorèmes du paragraphe précédent au triplet (X,M,D) , ce qui nous permettra d'introduire des notations utilisées fréquemment dans la suite de ce travail.

On a le triplet (X,M,D) ,où X est une matrice de données pxn , M une métrique sur $E = \mathbb{R}^p$,D une métrique diagonale sur $F = \mathbb{R}^n$; X est centrée pour

$$D = \begin{pmatrix} p_1 & & \\ & p_2 & 0 \\ & 0 & \ddots \\ & & & p_n \end{pmatrix} .$$



$W_X = X'MX$

$V_X = XDX'$

Avec les notations du théorème1 , on a $X = U\Delta V'$

- Les colonnes de U sont les vecteurs propres normés de V_XM , qui réalisent ,sous contrainte de M-orthonormalité, le maximum de $||\ X'Mu\ ||^2_D$.
- Soit $Y = \Delta V' = U'MX$; les lignes de Y sont les composantes principales, vecteurs propres de W_XD de norme égale à la racine carrée de la valeur propre correspondante .
- Soit $\Lambda = \Delta^2$, matrice diagonale des valeurs propres de W_XD ou de V_XM, $\lambda_1, \lambda_2, \ldots, \lambda_r$.

On a : $X = U\Delta V' = UY = \sum_{i=1}^{r} U_i Y^{i'}$ (Y^i est la $i^{ème}$ ligne de Y)

- Selon les théorèmes 2 et 3 , une approximation optimale de W_X de rang $s \leq r$,au sens de $||\ \ ||_{\varphi_{DD}}$ est fournie par :

$\sum_{i=1}^{s} Y_i Y_i'$; on représente les individus par les premières lignes de Y, en tenant compte, dans l'interprétation , des poids affectés aux individus .

Les lignes de $Y = U'MX$ sont les coordonnées des individus dans la base M-orthonormée des colonnes de U .Un individu supplémentaire ,vecteur x de E ,sera représenté par $U'Mx$.

- On représente de même les variables par les premières colonnes de $Z = U\Lambda^{1/2} = XDY'\Lambda^{-1/2}$, coordonnées des variables dans la base D-orthonormale des colonnes de $V = Y'\Lambda^{-1/2}$ .Une variable supplémentaire,y , vecteur de F , sera représentée par $y'DY'\Lambda^{-1/2}$.

- On est aussi dans le cas particulier du théorème 5 puisque D est diagonale , les colonnes de U réalisent donc le maximum de :

$\sum_{i=1}^{n} p_i < x_i, u >^2_M$,inertie du nuage des individus par rapport aux

espaces vectoriels orthogonaux à ceux engendrés successivement par les U_i ; les U_i engendrent les axes principaux d'inertie du nuage .

A part cette dernière propriété, tous les résultats d'Analyse en Composantes Principales sont conservés quand D n'est pas diagonale , mais est une métrique euclidienne quelconque sur F .

NOTATIONS :

Dans toute la suite on utilisera les notations ci-dessous:

- A partir du triplet (X_k, M_k, D_k) on construit :

$$W_k = X_k' M_k X_k \qquad \text{et} \qquad V_k = X_k D_k X_k'$$

- Λ_k est la matrice diagonale des valeurs propres non nulles de $W_k D_k$ ou de $V_k M_k$, rangées par ordre décroissant $\lambda_1, \lambda_2, \ldots, \lambda_{r_k}$.

- Y_k est la matrice $r_k x n_k$ dont les lignes sont des vecteurs propres de $W_k D_k$,associés aux λ_i ,de D_k-norme égale à $\sqrt{\lambda_i}$.

- Z_k est la matrice $p_k x r_k$ dont les colonnes sont des vecteurs propres de $V_k M_k$,associés aux λ_i ,de M_k-norme égale à $\sqrt{\lambda_i}$.

- U_k est défini par :

$$U_k = Z_k \Lambda_k^{-1/2}$$

L'indice k sera omis lorsque la matrice considérée est la même pour tout k .

*
* *

II : LES DONNEES - CARACTERISATION DES SITUATIONS RENCONTREES :

On dispose de K tableaux de données, quantitatifs ou qualitatifs, obtenus à différentes époques ou dans différentes occasions, et que l'on désire comparer.

Nous allons voir quelles sont les situations que l'on peut ainsi rencontrer, et comment dans chaque cas associer à chaque tableau un élément caractéristique.

1 / LES SITUATIONS :

Situation 1 : K matrices de données quantitatives concernant les mêmes individus, munis des mêmes poids. On a donc K triplets (X_k, M_k, D), où X_k est une matrice $p_k \times n$, centrée pour D, et M_k une métrique sur $\mathbb{R}^{p_k}$. On pose $W_k = X' M_k X$.

Situation 2 : K matrices de similarités, de dissimilarités, de distances entre les mêmes individus, munis des mêmes poids (matrice diagonale D). On substitue alors à chaque tableau une matrice W_k symétrique, semi-définie positive, centrée pour D, de façon à pouvoir comparer les différentes "visions" des n individus fournies par les K tableaux, au moyen des W_k.

- Si on a des matrices de similarités S_k, on construit $\hat{S}_k = (I - \underset{\sim}{1}\underset{\sim}{1}'D)\, S_k (I - D\underset{\sim}{1}\underset{\sim}{1}')$: Si $\hat{S}_k$ est semi-définie positive, on pose $W_k = \hat{S}_k$, sinon, soit $\hat{S}_k^* = \hat{S}_k + |\lambda_m| I$, où λ_m est la plus petite valeur propre de $\hat{S}_k$. On pose alors $W_k = (I - \underset{\sim}{1}\underset{\sim}{1}'D)\, \hat{S}_k^* (I - D\underset{\sim}{1}\underset{\sim}{1}')$. On montre que $\hat{S}_k^*$ est semi-définie positive, W_k a bien les propriétés cherchées (Voir par exemple (4) chapitre VIII ou (22)).

- Si on a des matrices de dissimilarités, de distances : Soit $\hat{D}_k$ la matrice dont les éléments sont les carrés des dissimilarités données ; on construit :

$\hat{S}_k = -\frac{1}{2}(I-\underset{\sim\sim}{1}1'D)\,\hat{D}_k(I-D\underset{\sim\sim}{1}1')$, et on procède ensuite comme précédemment pour obtenir W_k.

Pour la situation 2), on considèrera par la suite qu'on a uniquement des (W_k,D), où W_k est une matrice nxn symétrique, semi-définie positive, centrée pour D.

Situation 3 : K matrices quantitatives résultant de l'observation des mêmes variables. On a K triplets (X_k,M,D_k), où X_k est une matrice pxn_k, centrée pour D_k, et M une métrique sur $\mathbb{R}^p$.

Situation 4 : K matrices de variances-covariances entre les mêmes p variables. On a donc des (V_k, M), où V_k est une matrice pxp, symétrique, semi-définie positive, et M une métrique sur $\mathbb{R}^p$.

Situation 5 : K matrices de données quantitatives résultant de l'observation des mêmes variables sur les mêmes individus ; ce sont souvent des données chronologiques.
On est à la fois dans la situation 1) et 3).
On peut vouloir s'intéresser à l'évolution des positions relatives des individus, à l'évolution des positions relatives des variables, ou bien à l'évolution d'un individu ou d'une variable.
On a K triplets (X_k,M,D), X_k matrice pxn.

Situation 6 : K variables qualitatives mesurées sur les mêmes individus, munis des mêmes poids. On a alors des triplets $(X_k,D_{1/p_k},D)$, où X_k est la matrice des indicatrices des modalités de la kième variable, D_{p_k} la matrice diagonale des poids des différentes modalités ($D_{p_k} = X_k D X'_k$),

et $D_{1/p_k} = D_{p_k}^{-1}$.

On veut ici pouvoir représenter d'une part toutes les variables, d'autre part toutes les modalités des différentes variables.

Situation 7 : K tableaux de contingence, ou de probabilités, ou d'échanges, X_k, entre les deux mêmes variables qualitatives.

On considèrera alors plutôt les tableaux de Burt $B_k = \begin{pmatrix} D_p(k) & X_k \\ X_k' & D_q(k) \end{pmatrix}$, où $D_p(k)$ et $D_q(k)$ sont les matrices diagonales des poids marginaux déduits de X_k.

On va donc s'intéresser à des situations très variées ; mais nous allons voir que les méthodes existantes pour étudier ces situations sont souvent très semblables.

Remarque :

Lorsqu'on a un mélange de variables quantitatives et de variables qualitatives mesurées sur les mêmes individus, la pratique habituelle consiste à rendre qualitatives les variables quantitatives, et ainsi à se ramener à la situation 6.

2 / EQUIVALENCE ENTRE TABLEAUX DE DONNEES - ELEMENT CARACTERISTIQUE D'UNE ETUDE

Disposant de K tableaux de données du même type, on veut pouvoir les comparer. Pour cela, nous allons d'abord définir dans quelle mesure on peut considérer que deux tableaux sont équivalents, et donner un élément caractéristique de chaque classe d'équivalence. L'équivalence retenue dépend en grande partie de la situation considérée, mais aussi de l'optique de l'étude.

2.1 / Opérateurs

Si on est dans la situation 1, 2, ou 5, on veut comparer les liaisons entre les individus, d'un tableau à l'autre, c'est-à-dire comparer les matrices W_k. L'élément caractéristique d'une étude sera ici W_k :

Equivalence 1) : (X_1,M_1,D) et (X_2,M_2,D) sont équivalents si $W_1D = W_2D$.

Si on fait l'analyse en composantes principales des deux triplets, on aura la même représentaion des individus : $Y_1 = Y_2$, $\Lambda_1 = \Lambda_2$ (avec les notations de I-4).

Equivalence 1') : (X_1,M_1,D) et (X_2,M_2,D) sont équivalents si il existe $\alpha > 0$, tel que $W_1D = \alpha W_2D$.

Si on fait l'analyse en composantes principales des deux triplets, on aura des représentations des individus homothétiques : $Y_1 = \sqrt{\alpha}\, Y_2$, $\Lambda_1 = \alpha\Lambda_2$.

Proposition : Si $p_1 \geq p_2$, $(W_1D = \alpha W_2D) \Leftrightarrow (\exists H,\ p_1 x p_2 \ /\ H'M_1H = M_2,\ X_1 = \sqrt{\alpha}\, HX_2)$.

Démonstration :

* si $X_1 = \sqrt{\alpha}\, HX_2$, $X_1'M_1X_1 = \alpha X_2'H'M_1HX_2 = \alpha X_2'\, M_2\, X_2$: $W_1D = \alpha W_2D$

* si $W_1 = \alpha W_2$, $Y_1 = \sqrt{\alpha}\, Y_2$, $\Lambda_1 = \alpha\Lambda_2$

Soit $\tilde{U}_2$ une matrice $p_2 x p_2$, M_2-orthonormale, obtenue en complétant le système M_2-orthonormal des vecteurs colonnes de U_2.

Soit $\tilde{U}_1$ une matrice $p_1 x p_2$, M_1-orthonormale, obtenue de façon similaire.

Soit $\tilde{Y}$ la matrice $p_2 x n$: $\tilde{Y} = \begin{pmatrix} Y_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

On a $X_1 = U_1Y_1 = \sqrt{\alpha}\, U_1Y_2 = \sqrt{\alpha}\, \tilde{U}_1\tilde{Y}$ et $X_2 = U_2Y_2 = \tilde{U}_2\tilde{Y}$.

$X_1 = \sqrt{\alpha}\, (\tilde{U}_1\tilde{U}_2'M_2\tilde{U}_2)\tilde{Y} = \sqrt{\alpha}\, (\tilde{U}_1\tilde{U}_2'M_2)X_2$.

Soit $H = \tilde{U}_1\tilde{U}_2'M_2$, on a $H'M_1H = M_2\tilde{U}_2\tilde{U}_1'M_1\tilde{U}_1\tilde{U}_2'M_2 = M_2\tilde{U}_2\tilde{U}_2'M_2 = M_2$.

Si on se restreint aux métriques identité I, l'équivalence 1) signifie que les deux nuages de points se déduisent par des rotations et des symétries, l'équivalence 1') par des similitudes.

2.2 / Projecteurs

Dans le cas d'études portant sur les mêmes individus (situations 1, 2, 5, 6) on peut s'intéresser, non pas, comme ci-dessus, aux positions des individus les uns par rapport aux autres, mais plutôt au sous-espace vectoriel de $\mathbb{R}^n$, $\mathcal{E}_k$, engendré par les variables initiales. Cela revient à comparer les études au moyen des projecteurs D-orthogonaux sur $\mathcal{E}_k$: A_k.

Si X_k est de rang p_k, A_k s'écrit $X_k' (X_k D X_k')^{-1} X_k D$.
Avec les notations de (I,4), comme les composantes principales forment une base de $\mathcal{E}_k$, on a, quelque soit le rang de X_k, $A_k = Y_k' \Lambda_k^{-1} Y_k D$.

Equivalence 2) : (X_1,D) et (X_2,D) sont équivalents si $A_1 = A_2$.

Cette équivalence est indépendante de M_1 et M_2. Elle peut être aussi considérée dans la situation 2) : à partir de $W_k = Y_k' Y_k$, on construit :

$A_k = Y_k' \Lambda_k^{+1} Y_k D$.

Proposition : Si il existe M_1 et M_2 tels que (X_1,M_1,D) et (X_2,M_2,D) sont équivalents au sens 1 ou 1', ils le sont au sens 2).
Inversement, si (X_1,D) et (X_2,D) sont équivalents au sens 2), et si X_1 est de rang p_1 et X_2 de rang p_2 (ici $p_1 = p_2$), alors $(X_1,(X_1 D X_1')^{-1}, D)$ et $(X_2,(X_2 D X_2')^{-1}, D)$ sont équivalents au sens 1).

En effet : si $W_1D = \alpha W_2D$, $Y_1 = \sqrt{\alpha}\ Y_2$, $\Lambda_1 = \alpha\Lambda_2$, $A_1 = Y_1'\Lambda_1^{-1}Y_1D =$

$$\alpha Y_2' \frac{\Lambda_2^{-1}}{\alpha} Y_2D = A_2$$

si $A_1 = A_2$, $X_1'(X_1DX_1')^{-1}X_1D = X_2'(X_2DX_2')^{-1}X_2D$.

Si $A_1 = A_2$, $\Lambda_1^{-1/2}Y_1$ et $\Lambda_2^{-1/2}Y_2$ fournissent (en lignes) deux bases D-orthonormées du même sous-espace vectoriel de F, alors que si $W_1 = \alpha W_2$, ces bases sont égales.

Cas particulier : données qualitatives (situation 6)

Si X_k est le tableau des variables indicatrices des p_k modalités d'une variable qualitative, $\mathcal{E}_k$ est l'ensemble des variables quantitatives que l'on peut reconstruire à partir de la variable qualitative. Chaque élément de $\mathcal{E}_k$ est de la forme $X_k'(a)$, où a est un codage de la variable.

Si X_k n'est pas centrée, on a toujours $\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \underset{\sim}{1}$ dans $\mathcal{E}_k$ $\quad(\underset{\sim}{1} = X_k'\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix})$.

Soit Δ_1 la droite de F engendrée par $\underset{\sim}{1}$. $\mathcal{E}_k = \mathcal{E}_k^* \oplus \Delta_{\underset{\sim}{1}}$, où $\mathcal{E}_k^*$ est l'espace vectoriel engendré par les variables de $X_k^* = X_k(I-D\underset{\sim\sim}{1}\underset{\sim\sim}{1}')$, c'est-à-dire l'ensemble des variables quantitatives centrées que l'on peut reconstruire à partir de la variable qualitative.

Soit A_k le projecteur sur $\mathcal{E}_k$, et A_k^* le projecteur sur $\mathcal{E}_k^*$:
$(A_1 = A_2) \Longleftrightarrow (A_1^* = A_2^*)$.
Dans ce cas, le fait de centrer ou non les tableaux n'influe pas sur l'équivalence.
Comme ici on ne considère que les triplets :$(X_k, D_{1/p_k}, D)$ où $D_{1/p_k} = (X_kDX_k')^{-1}$, les équivalences 1), 1') et 2) coïncident.

2.3 / Matrices de variances-covariances

Dans le cas d'études résultant de l'observation des mêmes variables (situations 3, 4, 5), on prendra comme élément caractéristique la matrice V_k de variances-covariances : $V_k = X_k D_k X_k'$.

Equivalence 3) : (X_1, M, D_1) et (X_2, M, D_2) sont équivalents si $V_1 = V_2$.

Si on fait l'analyse en composantes principales des deux triplets, on aura la même représentation des variables : $Z_1 = Z_2$, $\Lambda_1 = \Lambda_2$ (avec les notations de I-4).

Equivalence 3') : (X_1, M, D_1) et (X_2, M, D_2) sont équivalents si il existe $\alpha > 0$ tel que $V_1 = \alpha V_2$.

Ici on aura des représentations homothétiques des variables :

$Z_1 = \sqrt{\alpha}\, Z_2$, $\Lambda_1 = \alpha \Lambda_2$.

2.4 / Tableaux de Burt

Dans le cas de la situation 7, on a des tableaux de contingence.

On sait que faire l'analyse des correspondances d'un tableau de contingence revient à faire l'analyse en composantes principales du tableau $\begin{pmatrix} X \\ Y \end{pmatrix}$ obtenu en juxtaposant les tableaux de variables indicatrices des modalités des deux variables qualitatives. La matrice de variances-covariances de cette étude est alors le tableau de Burt : $B_k = \begin{pmatrix} D_p(k) & X_k \\ X_k' & D_q(k) \end{pmatrix}$, où

$D_p(k) = XDX'$

$D_q(k) = YDY'$.

Equivalence 4) : $(X_1 = X_2) \iff (B_1 = B_2)$

Pour comparer les tableaux de contingence, on comparera les tableaux de Burt correspondants. C'est ce que fait T. FOUCARD (13).

2.5 / Cas de la situation 5

On a les mêmes variables observéessur les mêmes individus : on peut à la fois comparer les W_k, les A_k, les V_k. Selon l'optique de l'étude, on pourra être amené à préférer l'une ou l'autre des équivalences correspondantes, selon que l'on désire étudier l'évolution des liaisons entre les individus, ou entre les variables.

Proposition : Soit (X_1, M, D) et (X_2, M, D) deux études portant sur les mêmes individus et les mêmes variables. Si il existe $\alpha > 0$ tel que $X_1' M_1 X_1 = \alpha X_2' M_2 X_2$, et $\beta > 0$ tel que $X_1 D X_1' = \beta X_2 D X_2'$, alors : $\alpha = \beta$, $X_1 = \sqrt{\alpha}\, X_2$ (au signe près)

En effet : si $W_1 = W_2$, $Y_1 = \sqrt{\alpha}\, Y_2$ et $\Lambda_1 = \alpha \Lambda_2$

si $V_1 = \beta V_2$, $Z_1 = \sqrt{\beta}\, Z_2$ et $\Lambda_1 = \beta \Lambda_2$ $\Bigg\} \Rightarrow \alpha = \beta$

$$X_1 = U_1 Y_1 = Z_1 \Lambda_1^{-1/2} Y_1 = \sqrt{\beta}\, Z_2 \frac{1}{\sqrt{\beta}} \Lambda_2^{-1/2} \sqrt{\alpha}\, Y_2 = \sqrt{\alpha}\, Z_2\, \Lambda_2^{-1/2} Y_2$$

$$X_1 = \sqrt{\alpha}\, X_2.$$

Equivalence 5) : $X_1 = X_2$ (on a à la fois l'équivalence 1 et 3)

Equivalence 5') : $X_1 = \pm \sqrt{\alpha}\, X_2$ (on a à la fois l'équivalence 1' et 3').

Donc si l'on veut étudier à la fois les liaisons entre individus et les liaisons entre variables, on comparera directement les tableaux initiaux. C'est ce que fait P.A. JAFFRENOU (15).

BILAN

Situations et dimensions des tableaux	Eléments caractéristiques	Equivalences
1/ (X_k, M_k, D) $p_k$xn	W_k ou A_k	1) : $W_1 = W_2$ 1') : $W_1 = \alpha W_2$ ou 2) : $A_1 = A_2$
2/ (W_k, D) nxn	W_k ou A_k	1) , 1') , ou 2)
3/ (X_k, M_k, D) pxn$_k$	V_k	3) : $V_1 = V_2$ ou 3') : $V_1 = \alpha V_2$
4/ (V_k, M) pxp	V_k	3) ou 3')
5/ (X_k, M, D) pxn	W_k , A_k , V_k ou X_k	1) , 1') , 2) , 3), 3'), ou: 5) : $X_1 = X_2$ 5') : $X_1 = \alpha X_2$
6/ $(X_k, D_{1/p_k}, D)$ $p_k$xn variables indicatrices	$W_k = A_k$	2)
7/ X_k tableaux de contingence pxp	B_k	4) : $B_1 = B_2$

III : APPROCHE "INTERSTRUCTURE - COMPROMIS - INTRASTRUCTURE" :

On regroupe ici plusieurs techniques récemment développées en France pour l'étude des situations évoquées dans le chapitre précédent.

Ce sont :

- La méthode "STATIS" de H. L'HERMIER DES PLANTES et Y. ESCOUFIER (18) appliquée à diverses situations (19 et 20), et complétée par la méthode "STAVA" de M.C. PLACE (23).
- La méthode de T.FOUCART (13).
- La méthode de P.A.JAFFRENOU (15).

Le point commun de ces techniques est qu'elles sont constituées de trois étapes très semblables. Nous allons voir d'abord le schéma général de ces trois étapes, puis la façon dont chaque étape est conduite, selon la méthode considérée.

1 / SCHEMA GENERAL :

Représentation globale : On désire avoir une représentation globale de tous les tableaux. Pour cela, on va définir des distances entre éléments caractéristiques de chaque tableau, puis représenter graphiquement ces distances. Cette étape sera appelée, comme dans la méthode "STATIS" (18), l'étape de "l'interstructure".

Compromis : On cherche à résumer les K tableaux en un seul (ou en un petit nombre de tableaux) qui soit le plus représentatif de l'ensemble. On aura donc un compromis, optimal selon certain critère. Selon les types de situations considérées, et selon le critère retenu, on obtiendra plusieurs sortes de compromis.

Représentation détaillée : On cherche à représenter simultanément tous les individus, ou toutes les variables utilisées, de façon à pouvoir visualiser, et ceci est particulièrement utile lorsqu'on a des données chronologiques, l'évolution des variables ou des individus. Cette étape sera appelée, toujours par référence à la méthode "STATIS", l'étape des "intrastructures"

2 / INTERSTRUCTURE :

Pour comparer globalement les K tableaux, on construit une matrice $\mathcal{C}$, K x K , de produits scalaires entre objets caractéristiques des tableaux. Les vecteurs propres de $\mathcal{C}$ fournissent alors une représentation globale, analogue à une représentation d'Analyse en Composantes Principales. Deux cas seront alors à envisager, selon que l'on considère les études comme des "individus" ou comme des "variables" .

2.1 / LA MATRICE $\mathcal{C}$:

Les objets caractéristiques des études, définis au chapitre précédent, sont :

-soit des matrices W_k de produits scalaires entre individus
-soit des projecteurs A_k
-soit des matrices de variances-covariances V_k
-soit des tableaux de Burt B_k
-soit les tableaux initiaux X_k.

Dans tous les cas, ce sont des matrices d'applications linéaires d'un espace vectoriel euclidien dans un autre. On va pouvoir les comparer grâce au produit scalaire défini en I.2 . $\mathcal{C}$ sera la matrice K x K dont l'élément $\mathcal{C}_{kl}$ sera, selon les cas :

- $Tr(W_k D W_l D)$
- $Tr(A_k A_l)$
- $Tr(V_k M V_l M)$
- $Tr(B_k B_l)$
- $Tr(X_k D X_l' M)$

Plus précisément on a :

a) $Tr(W_k D W_l D) = Tr(X_k' M_k X_k D X_l' M_l X_l D)$

Proposition : $Tr(W_k D W_l D) \geq 0$
$(Tr(W_k D W_l D) = 0) \Leftrightarrow (X_k D X_l' = 0)$,dans ce cas les variables des deux tableaux sont non corrélées.

Démonstration : $M_k = L_k L_k'$ et $M_l = L_l L_l'$, où L_k (respectivement L_l) est une matrice carrée de rang p_k (respectivement p_l).

$$Tr(W_k D W_l D) = Tr(L_k' X_k D X_l' L_l L_l' X_l D X_k' L_k) = \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} (L_k' X_k D X_l' L_l)^2_{ij} \geq 0$$

et

$$Tr(W_k D W_l D) = 0 \iff L_k' X_k D X_l' L_l = 0 \iff X_k D X_l' = 0$$

Cas particuliers :

- Si $M_k = I_{p_k}$ et $M_l = I_{p_l}$, $Tr(W_k D W_l D) = \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} cov^2((X_k)^i, (X_l)^j)$

- Si $M_k = D_{1/\sigma_k^2}$ et $M_l = D_{1/\sigma_l^2}$, matrices diagonales des inverses des variances,

$$Tr(W_k D W_l D) = \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} cor^2((X_k)^i, (X_l)^j)$$

b) Si X_k est de rang p_k et X_l est de rang p_l, on a

$$Tr(A_k A_l) = Tr(X_k' (X_k D X_k')^{-1} X_k D X_l' (X_l D X_l')^{-1} X_l D)$$

sinon, on a toujours :

$$Tr(A_k A_l) = Tr(Y_k' \Lambda_k^{-1} Y_k D Y_l' \Lambda_l^{-1} Y_l D)$$

Proposition : // $Tr(A_k A_l) \geq 0$

// $Tr(A_k A_l) = 0 \iff X_k D X_l' = 0$

Démonstration : $Tr(A_k A_l) = Tr(\Lambda_k^{-1} Y_k D Y_l' \Lambda_l^{-1} Y_l D Y_k') = \sum_{i=1}^{r_k} \sum_{j=1}^{r_l} cor^2((Y_k)^i, (Y_l)^j)$, somme des carrés des coefficients de corrélation entre les composantes principales des deux études. Donc $Tr(A_k A_l) \geq 0$

et $Tr(A_k A_l) = 0 \iff Y_k D Y_l' = 0$.

Or comme $X_k = U_k Y_k$ et $X_l = U_l Y_l$,

$$Y_k D Y_l' = 0 \implies X_k D X_l' = U_k Y_k D Y_l' U_l' = 0,$$

et comme $Y_k = U_k' M_k X_k$ et $Y_l = U_l' M_l X_l$,

$$X_k DX_l' = 0 \Rightarrow Y_k DY_l' = U_k' M_k X_k DX_l' M_l U_l = 0$$

Donc

$$Y_k DY_l' = 0 \iff X_k DX_l' = 0$$

Ceci signifie que les sous-espaces vectoriels de F, $\mathcal{C}_k$ et $\mathcal{C}_l$, sont D-orthogonaux si et seulement si les variables des deux tableaux sont non corrélées.

Cas particulier : Situation 6 (variables qualitatives) :

$$Tr(A_k A_l) = Tr(X_k' D_{1/p_k} X_k DX_l' D_{1/p_l} X_l)$$

Si $D = \frac{1}{n} I_n$, $X_k DX_l'$ est la matrice P , d'élément $P_{ij} = \frac{n_{ij}}{n}$,

$$Tr(A_k A_l) = \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} \frac{P_{ij}^2}{P(i)P(j)} = \sum_{i=1}^{p_k} \sum_{j=1}^{p_l} \frac{(P_{ij}-P(i)P(j))^2}{P(i)P(j)} + 1$$

$$= \phi^2 + 1 = \frac{\chi^2}{n} + 1$$

Si on préfère considérer les tableaux centrés, et $A_k^* = A_k - A_{\Delta_j}$,

$$Tr(A_k^* A_l^*) = \phi^2 = \frac{\chi^2}{n} \quad .$$

c) $Tr(V_k M V_l M) = Tr(X_k D_k X_k' M X_l D_l X_l' M)$

On montre de façon analogue à ce qui a été fait en a) ceci :

Proposition : $Tr(V_k M V_l M) \geq 0$

$Tr(V_k M V_l M) = 0 \iff X_k' M X_l = 0$, dans ce cas les individus engendrent des sous-espaces vectoriels de E M-orthogonaux.

d) On a de même :

$Tr(B_k B_l) \geq 0$ si B_k et B_l sont des tableaux de Burt.

Proposition : La matrice $\mathcal{C}$ est semi-définie positive .

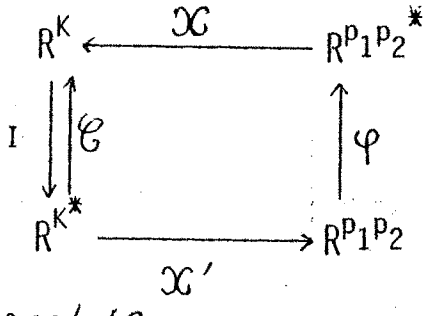
Démonstration : Soit O_k l'objet caractéristique de la $k^{ème}$ étude, et φ le produit scalaire entre objets($\mathcal{C}_{kl} = \langle O_k, O_l \rangle_\varphi$).

Soit $\alpha' = (\alpha_1 \ldots\ldots \alpha_k)$, $\alpha' \mathcal{C} \alpha = \sum_{k=1}^{K} \sum_{l=1}^{K} \alpha_k \alpha_l \langle O_k, O_l \rangle_\varphi = \| \sum_{k=1}^{K} \alpha_k O_k \|^2_\varphi \geq 0$

2.2 /DESCRIPTION DES ETUDES-"VARIABLES" :

Chaque O_k est une matrice $p_1 \times p_2$, à laquelle on peut associer le vecteur $\tilde{O}_k$ de $R^{p_1 p_2}$ obtenu en juxtaposant les colonnes de O_k. On note φ la métrique $M_2 \otimes M_1$, où M_1 est la métrique sur $E_1 = R^{p_1}$, et M_2 la métrique sur $E_2 = R^{p_2}$.

Soit $\mathcal{X}$ le tableau $K \times (p_1 p_2)$: $\mathcal{X} = \begin{pmatrix} \tilde{O}'_1 \\ \vdots \\ \tilde{O}'_k \end{pmatrix}$. On a le schéma :



On a $\mathcal{C} = \mathcal{X} \varphi \mathcal{X}'$. $\mathcal{C}$ est analogue à une matrice de variances covariances entre les K "variables" O_k . $\mathcal{C} = YY'$, où Y est la matrice $K \times r$ (r= rang de $\mathcal{C}$) dont les colonnes sont vecteurs propres de $\mathcal{C}$, de norme égale à la racine carrée de la valeur propre correspondante .

Les lignes de Y fournissent une représentation des K études analogue à celle des variables en Analyse en Composantes Principales; chaque étude est représentée par un vecteur qui, s'il est bien représenté, a pour norme $\| O_k \|_\varphi$; deux vecteurs bien représentés auront un angle proche de celui de O_k et O_l selon φ . En particulier, si O_k est de la forme W_k ou A_k , deux vecteurs seront orthogonaux si les variables de X_k et celles de X_l sont non corrélées. Deux études équivalentes au sens 1', 3' ou 5' seront représentées par des vecteurs colinéaires.

Si on veut représenter de façon identique des objets proportionnels (équivalence 1',3',5'), on diagonalisera plutôt la matrice $\mathcal{R}$ des cosinus entre objets :

$$\mathcal{R}_{kl} = \frac{< O_k, O_l >_\varphi}{\|O_k\|_\varphi \, \|O_l\|_\varphi}$$

Cela revient à normer toutes les"variables"O_k .

L'interprétation de la représentation obtenue est analogue à celle d'une Analyse en Composantes Principales sur matrice de corrélation ; une étude sera bien représentée si son point représentatif est proche du cercle de rayon unité. Les cosinus d'angles entre vecteurs associés à des études bien représentées sont alors proches des valeurs $\mathcal{R}_{kl}$ qui seront selon les cas considérés de la forme suivante :

- Si O_k est de la forme W_k ,

$$\mathcal{R}_{kl} = \frac{Tr(W_k D W_l D)}{\sqrt{Tr(W_k D)^2 Tr(W_l D)^2}} = Rv(W_k D, W_l D) .$$

- Si O_k est de la forme A_k ,

$$\mathcal{R}_{kl} = \frac{Tr(A_k A_l)}{\sqrt{Tr(A_k)^2 Tr(A_l)^2}} = \frac{Tr(A_k A_l)}{\sqrt{r_k r_l}}$$

où r_k et r_l sont les rangs de X_k et X_l .

Cas particulier : situation 6 (variables qualitatives) :

On a vu que :

$$Tr(A_k A_l) = \Phi^2 + 1 = \frac{\chi^2}{n} + 1 \quad \text{et} \quad Tr(A_k^* A_l^*) = \Phi^2 = \frac{\chi^2}{n}$$

On a aussi $Tr(A_k^2) = p_k$ et $Tr(A_k^{*2}) = p_k - 1$

donc $$\mathcal{R}_{kl} = \frac{\Phi^2 + 1}{\sqrt{p_k p_l}} \quad \text{ou} \quad \mathcal{R}_{kl} = \frac{\Phi^2}{\sqrt{(p_k - 1)(p_l - 1)}} ,$$

cette dernière expression étant le coefficient de Tchuprov,noté T^2 , permettant de mesurer le degré d'association entre deux variables qualitatives,qui est nul lorsque les deux variables qualitatives sont indépendantes.

- Si 0_k est de la forme V_k ,

$$\mathcal{R}_{kl} = \frac{Tr(V_k M V_l M)}{\sqrt{Tr(V_k M)^2 Tr(V_l M)^2}}$$

- Si 0_k est de la forme X_k ,

$$\mathcal{R}_{kl} = \frac{Tr(X_k D X_l' M)}{\sqrt{Tr(X_k D X_k' M)^2 Tr(X_l D X_l' M)^2}} = \frac{Tr(X_k D X_l' M)}{\sqrt{In(X_k)\ In(X_l)}} ,$$

où In(X) représente l'inertie du nuage des individus par rapport au centre de gravité.

Remarques : 1/ Si 0_k est de la forme W_k, $\Psi = D \otimes D$; si 0_k est de la forme A_k, $\Psi = I$; Ψ est donc une matrice diagonale de poids, et chaque "variable" 0_k est centrée pour Ψ (si X_k est centrée):

$$\sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j (W_k)_{ij} = 0 .$$

$\mathcal{C}$ est véritablement une matrice de variances-covariances et $\mathcal{R}$ une matrice de corrélation. Il en est de même si 0_k est de la forme X_k et si M = I (Alors $\Psi = D \otimes I$).

2/ Si on est dans la situation 5 (mêmes variables, mêmes individus), on peut envisager de comparer les études de trois façons:

a/ à partir de $\mathcal{C}_1$: $(\mathcal{C}_1)_{kl} = Tr(W_k D W_l D)$,

b/ à partir de $\mathcal{C}_2$: $(\mathcal{C}_2)_{kl} = Tr(V_k M V_l M)$,

c/ à partir de $\mathcal{C}_3$: $(\mathcal{C}_3)_{kl} = Tr(X_k D X_l' M)$.

Les trois représentations obtenues par diagonalisation des $\mathcal{C}_i$ peuvent être très différentes, car elles mettent en évidence des liens de nature différente entre les études. On peut seulement remarquer que:

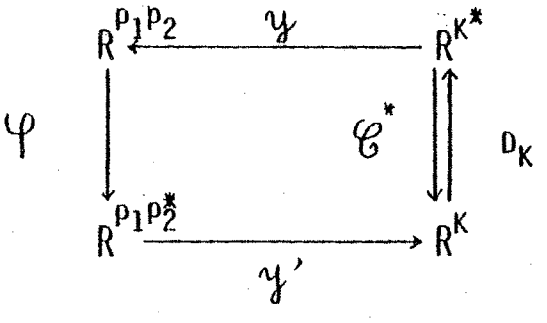
- Deux études sont colinéaires pour c/ si et seulement si elles le sont pour a/ et pour b/ .
- Si deux études sont orthogonales pour a/ ($X_k D X_l' = 0$) ou pour b/ ($X_k' M X_l = 0$), elles le sont pour c/.
- $\mathcal{C}_1$ et $\mathcal{C}_2$ ont la même diagonale : $Tr(W_k D)^2 = Tr(X_k' M X_k D X_k' M X_k D)$
$= Tr(X_k D X_k' M X_k D X_k' M) = Tr(V_k M)^2$.

2.3 / DESCRIPTION DES ETUDES-"INDIVIDUS":

La matrice $\mathcal{C}$,symétrique et semi-définie positive,peut être considérée comme une matrice de produits scalaires entre K individus. Cependant,pour pouvoir l'assimiler à une matrice du type "W" du schéma de dualité,on la centre par rapport à la matrice diagonale des poids,D_K,que l'on peut affecter a priori aux études.On remplace $\mathcal{C}$ par:

$\mathcal{C}^* = (I - \underset{\sim}{1}\underset{\sim}{1}'D_K)\,\mathcal{C}\,(I - D_K\underset{\sim}{1}\underset{\sim}{1}')$,où $\underset{\sim}{1}$ est le vecteur $\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}$ de R^K.

$\mathcal{C}^*$ engendre les mêmes distances entre études que $\mathcal{C}$,mais fournit une représentation centrée.On a le schéma:



$\mathcal{C}^* = \mathcal{Y}'\varphi\mathcal{Y}$,où $\mathcal{Y} = \mathcal{X}'(I - D_K\underset{\sim}{1}\underset{\sim}{1}') = (\tilde{0}_1 \ldots \tilde{0}_K)(I - D_K\underset{\sim}{1}\underset{\sim}{1}')$ $_K$

Centrer $\mathcal{C}$ revient à remplacer chaque objet 0_k par $0_k - \sum_{k=1} q_k 0_k$,où q_k est le poids affecté à la $k^{ième}$ étude.

Les premiers vecteurs propres de $\mathcal{C}D_k$ de norme égale à la racine carrée de la valeur propre correspondante,fournissent une représentation des K études telle que les distances entre points représentatifs des 0_k soient les meilleures approximations possibles des distances $d^2(0_k,0_l) = ||0_k - 0_l||^2_\varphi$.

3 / COMPROMIS :

On veut résumer les K objets $0_1,0_2,\ldots,0_K$ par un seul,considéré comme "représentatif"de l'ensemble.

Compromis n°1 :

Si $q_1,q_2,\ldots,q_K$ sont des poids affectés a priori aux différentes

études,on peut choisir comme compromis $\bar{0} = \sum_{k=1}^{K} q_k 0_k$; c'est l'étude-"individu" par rapport à laquelle l'inertie du nuage des études-"individus" est la plus faible:

Propriété : $\bar{0}$ réalise le minimum de $\sum_{k=1}^{K} q_k ||0_k - \sum_{l=1}^{K} \alpha_l 0_l||_\varphi^2$.

En effet: $$\sum_{k=1}^{K} q_k ||0_k - \sum_{l=1}^{K} \alpha_l 0_l||_\varphi^2 = \sum_{k=1}^{K} q_k \left(||0_k - \bar{0}||_\varphi^2 + ||\bar{0} - \sum_{l=1}^{K} \alpha_l 0_l||_\varphi^2 \right)$$

est minimum pour $\alpha_l = q_l$ pour tout l.
En général,on a $q_k = \frac{1}{K}$ pour tout k,donc $\bar{0} = \frac{1}{K} \sum_{k=1}^{K} 0_k$:on prend comme objet compromis la moyenne des objets caractéristiques de toutes les études .

Compromis $\overset{\circ}{n=2}$:

Un autre compromis possible est donné par $0 = \sum_{k=1}^{K} l_k 0_k$,où $l = \begin{pmatrix} l_1 \\ \vdots \\ l_K \end{pmatrix}$ est un vecteur propre de $\mathcal{C}$,normé,associé à la plus grande valeur propre, λ_1 ; 0 est la première composante principale de l'analyse de $(\mathcal{X}, I, \varphi)$de 2-2 .

Propriétés : Pour tout vecteur $\alpha' = (\alpha_1\ \alpha_2 \ldots \alpha_K)$ tel que $\alpha'\alpha = 1$, on a :

1/ $||\sum_{k=1}^{K} \alpha_k 0_k||_\varphi^2 \leq ||0||_\varphi^2 = \lambda_1$

2/ $\sum_{k=1}^{K} < \sum_{k=1}^{K} \alpha_k 0_k, 0_k >_\varphi^2 \leqslant \sum_{k=1}^{K} < 0, 0_k >_\varphi^2 = \lambda_1^2$.

Démonstration :

Ces propriétés sont l'application à $(\mathcal{X}, I, \varphi)$ des résultats du théorème 5 de I-3,et de la remarque qui suit ce théorème:
l réalise le maximum de $||\mathcal{X}' I u||_\varphi^2$,qui vaut λ_1.
$\frac{1}{\sqrt{\lambda_1}}$ 0 réalise le maximum de $\sum_{k=1}^{K} < \sum_{k=1}^{K} \beta_k 0_k, 0_k >_\varphi^2$,avec $\sum_{k=1}^{K} \beta_k^2 = \frac{1}{\sqrt{\lambda_1}}$,et le maximum vaut λ_1, donc 0 réalise le maximum de $\sum_{k=1}^{K} < \sum_{k=1}^{K} \alpha_k 0_k, 0_k >_\varphi^2$ avec $\sum_{k=1}^{K} \alpha_k^2 = 1$, et ce maximum vaut λ_1^2 .

Si on désire avoir plusieurs compromis,on peut prendre d'autres composantes principales de $(\mathcal{X}, I, \varphi)$,mais les compromis obtenus n'ont pas toujours un sens .

En effet,il est préférable d'avoir,comme compromis des O_k,un objet du même type qui puisse être considéré comme l'objet caractéristique d'une étude.

- Cas des W_k:

Comme tous les éléments de $\mathcal{C}$ sont positifs,les composantes de 1 sont toutes de même signe(théorème de Frobenius);on peut les choisir toutes positives.O (ainsi que $\bar{O}$) est une combinaison linéaire à coefficients positifs de matrices semi-définies positives,les deux compromis sont semi-définis positifs,et centrés pour D,on a bien des compromis du type "W".

De plus,àchaque choix de coefficients $\alpha_1,\alpha_2,\ldots\alpha_K$ positifs,on peut faire correspondre,si on a au départ des études (X_k,M_k,D),le schéma de dualité:

$$\begin{array}{ccc} R^{\Sigma p_k} = E & \xleftarrow{\quad X \quad} & F^* \\ \downarrow M & & \uparrow W \\ E^* & \xrightarrow[X']{} & F = R^n \end{array} \qquad \text{(D : } F \to F^*\text{)}$$

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} \qquad M = \begin{pmatrix} \alpha_1 M_1 & & & 0 \\ & \alpha_2 M_2 & & \\ & & \ddots & \\ 0 & & & \alpha_K M_K \end{pmatrix}$$

$$W = X'MX = \sum_{k=1}^{K} \alpha_k X_k' M_k X_k = \sum_{k=1}^{K} \alpha_k W_k \ .$$

La deuxième composante principale de $(\mathcal{C}, I, \varphi)$ est une combinaison linéaire des W_k dont certains coefficients sont forcément négatifs.Elle n'est donc pas toujours semi-définie positive.Ce "compromis" n'aurait alors pas beaucoup de sens ici.

- Cas des A_k :

En général,une combinaison linéaire de projecteurs n'est pas un projecteur.Les compromis obtenus ne sont donc pas vraiment de type "A",mais seulement "W".

En Analyse Canonique Généralisée (références (5) ou(16)),on est conduit,pour caractériser les positions relatives des K espaces vectoriels $\mathcal{E}_k$ engendrés par les lignes des X_k,à diagonaliser $\sum_{k=1}^{K} A_k$.
Ceci généralise l'analyse canonique de deux tableaux qui peut être considérée comme l'Analyse en Composantes Principales du triplet:

$$\left(\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \begin{pmatrix} (X_1 D X_1')^{-1} & 0 \\ 0 & (X_2 D X_2')^{-1} \end{pmatrix}, D\right) .$$

Les vecteurs propres normés de $\sum_{k=1}^{K} A_k$ réalisent successivement le maximum de $\sum_{k=1}^{K} ||A_k x||_D^2$,pour x'Dx=1,sous contrainte d'orthogonalité entre les x .
Si x est un vecteur propre de $\sum_{k=1}^{K} A_k$,associé à la valeur propre k,il se trouve dans l'intersection de tous les $\mathcal{E}_k$.

Cette technique,appliquée à la situation 6,donne l'Analyse des Correspondances Multiples,permettant de décrire les liens entre K variables qualitatives mesurées sur les mêmes individus.
On peut,pour cette même situation,utiliser le compromis n°2,qui a d'autres propriétés optimales.Cela revient à effectuer l'Analyse des Correspondances du tableau $\begin{pmatrix} \sqrt{\alpha_1}\, X_1 \\ \vdots \\ \sqrt{\alpha_K}\, X_K \end{pmatrix}$ (Voir (8)).

- Cas des X_k :

P.A JAFFRENOU (15) étudie le cas où on considère comme "compromis" toutes les composantes principales de $(\mathcal{X}, I, \varphi)$;ce sont des éléments de $\mathcal{M}_{pxn}$,donc des tableaux du même type que les X_k,centrés pour D .

Ces compromis forment une base de $\mathcal{L}'(\mathbb{R}^{K^*})$, dans laquelle on peut représenter tous les tableaux initiaux .

- Cas des B_k :

Si $B_1, B_2, \ldots, B_K$ sont des tableaux de Burt, $\bar{B} = \frac{1}{K} \sum_{k=1}^{K} B_k$ en est un. (Cela revient à faire la moyenne des tableaux de contingence).

Mais $\sum_{k=1}^{K} l_k B_k$ n'est pas un tableau de Burt.

T.FOUCART (13) propose comme compromis $B = \frac{1}{\sum_{k=1}^{K} l_k} \sum_{k=1}^{K} l_k B_k$. Ce compromis n'a pas les propriétés optimales des compromis déjà vus, mais il revient à prendre le compromis n=1, en affectant à chaque étude un poids égal à

$q_k = \frac{l_k}{\sum_{k=1}^{K} l_k}$, d'autant plus faible que l'étude considérée est loin des autres.

Remarque :

Dans le cas des B_k on est amené à se restreindre à des compromis $\sum_{k=1}^{K} \alpha_k O_k$, tels que $\alpha_1 + \alpha_2 + \ldots + \alpha_K = 1$, alors que pour les propriétés optimales du compromis n°2 , on a la contrainte : $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_K^2 = 1$.

Proposition :

Le maximum de $\| \sum_{k=1}^{K} \alpha_k O_k \|_\varphi^2$, avec $\alpha_k \geq 0$ et $\sum_{k=1}^{K} \alpha_k = 1$, est atteint par O_k tel que $\| O_k \|_\varphi^2 = \max \left\{ \| O_l \|_\varphi^2, l=1,\ldots,K \right\}$.

Démonstration :

$$\left\| \sum_{k=1}^{K} \alpha_k O_k \right\|_\varphi^2 = \sum_{k=1}^{K} \sum_{l=1}^{K} \alpha_k \alpha_l \langle O_k, O_l \rangle_\varphi \leq \sum_{k=1}^{K} \sum_{l=1}^{K} \alpha_k \alpha_l \| O_k \|_\varphi \| O_l \|_\varphi$$

$$\leq \left(\max_{k=1,\ldots,K} \| O_k \|_\varphi^2 \right) \sum_{k=1}^{K} \sum_{l=1}^{K} \alpha_k \alpha_l$$

$$= \left(\max_{k=1,\ldots,K} \| O_k \|_\varphi^2 \right) \left(\sum_{k=1}^{K} \alpha_k \right)^2 = \max_{k=1,\ldots,K} \| O_k \|_\varphi^2 .$$

Donc si on veut conserver les propriétés optimales du compromis $\overset{\circ}{n}=2$, avec la contrainte $\alpha_1+\alpha_2+\ldots+\alpha_K=1$,il faut prendre comme compromis l'objet de norme maximale .La méthode proposée par T. FOUCART semble préférable à cette solution extrême .

- Pratique habituelle dans le cas de la situation 5 :

Disposant de K tableaux de données concernant les mêmes variables et les mêmes individus,de nombreux auteurs se contentent de réaliser l'Analyse en Composantes Principales,

-Soit de $X=\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix}$,en diagonalisant $W_X D = \sum_{k=1}^{K} W_k D$,

ou $V_X = \begin{pmatrix} V_1 & X_1 D X_2' & \ldots & X_1 D X_K' \\ \vdots & & & \vdots \\ X_K D X_1' & \ldots\ldots & & V_K \end{pmatrix}$,

ce qui revient à diagonaliser,pour étudier l'ensemble des variables,le compromis "moyen" ($\overset{\circ}{n}=1$) entre les W_k ,

-Soit de $\mathcal{Z} = (X_1\ X_2\ \ldots X_K)$,en diagonalisant $V_{\mathcal{Z}} = \sum_{k=1}^{K} V_k$,

ou $W_{\mathcal{Z}} \begin{pmatrix} D & & 0 \\ & \ddots & \\ 0 & & D \end{pmatrix} = \begin{pmatrix} W_1 D & X_1' X_2 D & \ldots & X_1' X_K D \\ \vdots & & & \vdots \\ X_K' X_1 D & \ldots\ldots\ldots & & W_K D \end{pmatrix}$,

ce qui revient,pour étudier l'ensemble des individus,à diagonaliser (à $\frac{1}{K}$ près) le compromis "moyen" entre les matrices de variances-covariances V_k ,

-Soit de $\overline{X} = \frac{1}{K} \sum_{k=1}^{K} X_k$,compromis "moyen" entre les tableaux initiaux X_k,

ce qui conduit à diagonaliser $WD = \frac{1}{K^2} \sum_{k=1}^{K} \sum_{l=1}^{K} (X_k' X_l D)$

ou $V = \frac{1}{K^2} \sum_{k=1}^{K} \sum_{l=1}^{K} (X_k D X_l')$.

L'exposé ci-dessus au sujet des compromis et de leurs propriétés a permis d'apporter quelques justifications à ces pratiques,et de proposer des

pratiques optimales dans un autre sens.

- Liens avec l'"Interstructure" :

Pour visualiser les positions des études par rapport aux compromis,on peut projeter ceux-ci dans le plan principal de l'Interstructure:

- $\overline{0}$ sera projeté au barycentre des points représentatifs des études,affectés des poids q_k . Dans le cas des études-"individus" , $\overline{0}$ se trouve au centre,en $\underline{0}$,et la norme du vecteur joignant $\underline{0}$ au point représentatif d'un individu 0_k est une approximation de la distance entre $\overline{0}$ et 0_k .

- 0,première composante principale,est projetée sur le premier axe,à une distance $\sqrt{\lambda_1}$. Le premier axe de l'Interstructure permet de voir quels sont les tableaux qui sont le plus "corrélés" (au sens φ) avec 0 .

4 / INTRASTRUCTURE :

La diagonalisation du compromis nous donne,dans le cas des W_k et des A_k , une représentation des individus "intermédiaire" entre les représentations que pourrait nous fournir chaque étude ; de même,dans le cas des V_k, on a une représentation "moyenne" des variables .
Dans le cas des B_k ,l'Analyse des Correspondances de B donne une représentation "moyenne" des modalités des variables .
On aimerait,en plus de ces résultats ,pouvoir représenter dans un même plan tous les individus,ou toutes les variables,ou les modalités de toutes les variables qualitatives . Nous allons voir comment résoudre ce problème .

4-1 / CAS OU ON A LES MEMES INDIVIDUS :

On est dans la situation 1 , 2 , 5 , ou 6 .0_k est de la forme W_k ,ou A_k . Tout ce qui suit sera développé dans le cadre de la situ-

ation 1 ,avec $O_k = W_k$, $W = \sum_{k=1}^{K} \alpha_k W_k$,mais on verra que c'est encore valable pour les autres situations,et pour $O_k = A_k$.

On a vu que l'étude du compromis revenait à considérer le schéma de dualité:

$$\begin{array}{ccc} R^{k=1}{}^{K}{}^{\Sigma} p_k = E & \xleftarrow{\quad X \quad} & F^* \\ M \downarrow & & W \quad \| D \\ E^* & \xrightarrow[X']{} & F = R^n \end{array}$$

$X = \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix}$, $M = \begin{pmatrix} \alpha_1 M_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_K M_K \end{pmatrix}$, les α_k étant des nombres positifs.

4-1-1 / Représentation des variables :

Si on effectue l'Analyse en Composantes Principales du triplet (X,M,D),une représentation des variables est donnée naturellement par $\hat{Z} = XDY'\Lambda^{-1/2}$,où Y est la matrice contenant en lignes les composantes principales,et $\Lambda = YDY'$.

$\hat{Z}_k = X_k DY'\Lambda^{-1/2}$ représente les variables de la $k^{ième}$ étude,dans la base D-orthonormale des lignes de $\Lambda^{-1/2}$ Y .
Dans cette même base,on peut représenter les composantes principales de chaque étude par:

$$\tilde{Y}_k = Y_k DY'\Lambda^{-1/2}$$

(On peut en effet considérer que $W = \sum_{k=1}^{K} \alpha_k W_k = \sum_{k=1}^{K} \alpha_k Y_k' Y_k$,on a les mêmes Y et Λ que pour l'analyse de $\left(\begin{pmatrix} Y_1 \\ \vdots \\ Y_K \end{pmatrix}, \begin{pmatrix} \alpha_1 I & & \\ & \ddots & \\ & & \alpha_K I \end{pmatrix}, D \right)$).

Si on garde Y en entier,on a ainsi une représentation exacte de toutes les variables et de toutes les composantes principales, les covariances entre variables étant conservées:

$$\hat{Z}_k D \hat{Z}_k' = V_k \qquad \text{et} \qquad Y_k D Y_k' = \Lambda_k$$

Si on ne garde que les deux premières lignes de Y ,pour avoir une représentation plane,on a la meilleure représentation plane simultanée de toutes les variables.

Remarque : Cette représentation plane est celle utilisée dans la méthode STATIS (18) ,mais elle y est présentée uniquement comme la représentation de points supplémentaires.En fait,on a montré ici que l'on a plus: quand on prend Y en entier,la représentation est exacte .

Dans le cas de la situation 6 (X_k tableau de variables indicatrices),on peut ainsi représenter simultanément toutes les modalités des variables qualitatives considérées .

4-1-2 / REPRESENTATION DES INDIVIDUS :

On voudrait pouvoir représenter simultanément les nK individus,vus par les K études,de façon à pouvoir suivre l'évolution des individus au travers des études.

On va voir quelles méthodes on été proposées jusqu'à présent,et en présenter une autre qui n'a pas les mêmes propriétés.

1ière méthode : M.C. PLACE utilise dans sa thèse (23),pour représenter les individus,les colonnes de :

$$\hat{Y}_k = Y_k D Y' (Y D Y')^{-1} Y = Y_k D Y' \Lambda^{-1} Y .$$

Cela revient à projeter toutes les composantes principales sur le même espace vectoriel ,celui engendré par les lignes de Y .
La proposition ci-dessous peut permettre de mieux comprendre ce que signifient les $\hat{Y}_k$:

Proposition : 1/ Si on garde pour Y toutes les composantes de WD ,

$$\hat{Y}_k = Y_k$$

2/ Si on ne garde pour Y que les s premières composantes de WD (s inférieur ou égal au rang de WD), la matrice:

$\begin{pmatrix} \hat{Y}_1 \\ Y_2 \\ \vdots \\ Y_K \end{pmatrix}$ réalise le minimum, parmi les matrices $B = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_K \end{pmatrix}$

de rang s, de $\sum_{k=1}^{K} \alpha_k \, ||Y_k - B_k||^2_{\varphi_{M_k} D}$.

(Les α_k sont les coefficients du compromis choisi: $W = \alpha_1 W_1 + \ldots + \alpha_K W_K$) .

Démonstration : Cette proposition n'est que l'application des théorèmes 1 et 4 du chapitre I au triplet $(\mathcal{Y}, M, D)$, où M est défini comme ci-dessus et $\mathcal{Y}' = (Y_1' \; Y_2' \; \ldots \; Y_K')$.

En effet : $\mathcal{Y}' M \mathcal{Y} = Y'Y$ et $YDY' = \Lambda$.

1/ Selon le théorème de décomposition, on a :

$\mathcal{Y} = \hat{Z} \, \tilde{\Lambda}^{1/2} Y = \mathcal{Y} DY' \Lambda^{-1} Y$,donc $Y_k = Y_k DY' \Lambda^{-1} Y = \hat{Y}_k$ pour tout k .

2/ Soit Y_o la matrice des s premières lignes de Y, et $\Lambda_o = Y_o DY_o'$ Selon le théorème 4, la matrice $\hat{B} = \mathcal{Y} \, DY_o' \Lambda_o^{-1} Y_o$ réalise le minimum de $||\mathcal{Y} - B||^2_{\varphi_{MD}}$ parmi les applications de rang inférieur ou égal à s, et

$$||\mathcal{Y} - \hat{B}||^2_{\varphi_{MD}} = \mathrm{Tr}\left((\mathcal{Y} - \hat{B})' M (\mathcal{Y} - B) D \right) = \sum_{k=1}^{K} \alpha_k \, \mathrm{Tr}((Y_k - B_k)' M_k (Y_k - B_k) D)$$

$$= \sum_{k=1}^{K} \alpha_k \, ||Y_k - B_k||^2_{\varphi_{M_k} D} \quad \text{,avec } B_k = Y_k DY_o' \, \Lambda_o^{-1} Y_o = \hat{Y}_k \text{ obtenue}$$

en ne gardant que les s premières lignes de Y .

Avec cette méthode il semble difficile de donner une interprétation de l'évolution des individus,d'une étude à l'autre,ainsi que des liens entre individus et variables au travers d'une représentation conjointe .

2ième méthode : C'est celle utilisée par H.L'HERMIER DES PLANTES pour STATIS. On représente les individus par les colonnes de :

$$\tilde{\hat{Y}}_k = (\Lambda^{-1/2} Y D Y_k' \Lambda_k^{-1/2}) Y_k \quad .$$

Ceci est justifié par le fait que les lignes de $\Lambda^{-1/2}Y$ forment une base D-orthonormée du sous-espace vectoriel de $\mathbb{R}^n$ engendré par les lignes de $\mathcal{Y}$, c'est-à-dire par les lignes de tous les Y_k .

$\tilde{\Lambda}^{1/2} Y D Y_k' \Lambda_k^{-1/2}$ représente les cosinus entre les variables du tableau Y_k et celles de Y . $\tilde{\hat{Y}}_k$ représente donc les variables de Y_k dans un espace qui englobe toutes les composantes principales de toutes les études.

Nous verrons au chapitre V (Analyse Procustéenne) une autre justification de cette méthode.

Si on représente conjointement les individus et les variables (composantes principales) de Y_k, on obtient, pour chaque tableau, les mêmes positions relatives des variables et des individus que celles de la représentation conjointe résultant de l'Analyse en Composantes Principales de (X_k, M_k, D)., à condition de garder toutes les composantes de Y .

Remarques : 1/ Si Y est pris en entier,

$$\tilde{\hat{Y}}_k' \tilde{\hat{Y}}_k = Y_k' \Lambda_k^{-1/2} Y_k D A_Y Y_k' \tilde{\Lambda}_k^{1/2} Y_k = Y_k' \Lambda_k^{-1/2} Y_k D Y_k' \Lambda_k^{-1/2} Y_k$$

$$= Y_k' Y_k = W_k$$

Les produits scalaires, et donc les distances, entre individus d'un même tableau sont exactement représentés.

2/ Si on est dans la situation où $D_k = A_k$, on a: $\tilde{\hat{Y}}_k = \tilde{\Lambda}^{1/2} Y A_k'$; les lignes de $\tilde{\hat{Y}}_k$ sont les projections sur $\mathcal{E}_k$ des vecteurs de la base orthonormée $Y'\Lambda^{-1/2}$ (vecteurs colonnes), dans ce cas,

$$\sum_{k=1}^{K} \alpha_k \tilde{\hat{Y}}_k = \tilde{\Lambda}^{1/2} Y \left(\sum_{k=1}^{K} \alpha_k A_k' \right) = \Lambda^{-1/2} Y D Y' Y = \Lambda^{1/2} Y$$

Aux λ_i près, l'individu "compromis" est au barycentre des représentations de cet individu au travers des K études, affectées des poids α_k .

3ième méthode : On peut représenter les individus par les colonnes de:

$$\hat{X}_k = \Lambda^{-1/2} YDW_k$$

Dans le cas où $O_k = A_k$, $\hat{X}_k = \hat{\hat{Y}}_k$,cette méthode revient à la précédente . Sinon on a dans tous les cas :

Propriétés : 1/ $\sum_{k=1}^{K} \alpha_k \hat{X}_k = \Lambda^{1/2} Y$

2/ $Tr\ (\hat{X}_k' \hat{X}_k D) = Tr\ (W_k D\ Y' \Lambda^{-1} YDW_k D) = Tr\ (W_k D W_k D)$

L'inertie du nuage représentant les individus de X_k est égale au carré de la norme de l'opérateur correspondant . On a de même:

3/ $Tr\ (\hat{X}_k D \hat{X}_l') = Tr\ (W_k D W_l D)$

Les distances entre $\hat{X}_k$ et $\hat{X}_l$ au sens φ_{ID} sont égales aux distances entre opérateurs correspondants pour φ_{DD},c'est-à-dire qu'on peut reconstituer la distance entre tableaux à partir de la somme pondérée(par D)des distances au carré entre les points homologues de $\hat{X}_k$ et de $\hat{X}_l$.

Remarques: 1/On peut aussi,en tenant compte du fait que la représentation des individus "compromis" et celle des variables résultent de l'Analyse en Composantes Principales de (X,M,D),projeter les individus de X_k en "individus supplémentaires",en assimilant le tableau X_k à

$\begin{pmatrix} 0 \\ \vdots \\ X_k \\ \vdots \\ 0 \end{pmatrix}$,de dimensions $(\sum_{k=1}^{K} p_k) \times n$. On a alors:

$$\hat{X}_k = \Lambda^{-1/2} (\hat{Z}_1' \ \ldots \ \hat{Z}_K') M \begin{pmatrix} 0 \\ \vdots \\ X_k \\ \vdots \\ 0 \end{pmatrix} = \alpha_k \Lambda^{-1/2} \hat{Z}_k' M_k X_k = \alpha_k \Lambda^{-1} YDW_k .$$

Ici $\sum_{k=1}^{K} \hat{X}_k = Y$, mais on n'a plus les propriétés 2/ et 3/ ci-dessus .

2/ $W_k D$ étant l'opérateur de $\mathbb{R}^n$ qui à toute variable y fait correspondre $\sum_{j=1}^{p_k} (X_k)^j \ cov(\ (X_k)^j, y), X_k'$ correspond aux valeurs de cet opérateur sur les variables de $Y' \Lambda^{-1/2}$.

Conclusion : On peut envisager différentes méthodes de représentation de tous les individus ,lorsqu'on dispose d'études portant sur les mêmes individus . Cependant,lorsque les variables diffèrent d'une étude à l'autre,il paraît difficile de comparer,pour un individu donné, les différentes valeurs que prennent les variables pour cet individu . On ne peut envisager que d'étudier l'évolution,non pas de l'individu lui-même,mais de ses liaisons avec les autres individus . C'est pourquoi il semble préférable d'utiliser la $3^{ième}$ méthode,qui ne fait intervenir que les W_k,et pas les X_k ni les Y_k .

4-2 / CAS OU ON A LES MEMES VARIABLES :

On est dans la situation 3 , 4 , ou 5 . C'est le cas dual du précédent : ici on considère $(X_1\ X_2\ \dots\ X_K)$,avec $\begin{pmatrix} D_1 & & 0 \\ & D_2 & \\ 0 & & \ddots \\ & & & D_K \end{pmatrix}$ comme matrice des poids .

-Les individus sont représentés par :

$\mathcal{Y}_k = \Delta^{-1/2} Z'MX_k$,où Z est tel que $V = \sum_{k=1}^{K} V_k = Z\,Z'$ et $Z'MZ = \Delta$.

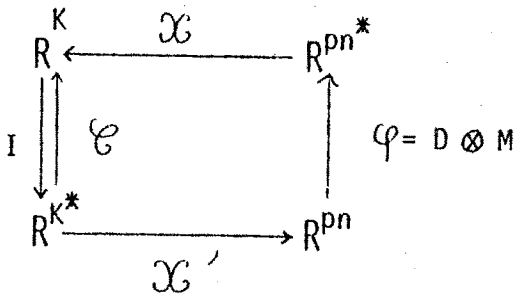
-Pour la représentation des variables,on a les mêmes difficultés que pour la représentation des individus dans le cas précédent. On peut utiliser les trois méthodes duales des précédentes .

4-3 /CAS OU ON A LES MEMES INDIVIDUS ET LES MEMES VARIABLES:

On est dans la situation 5 ou 7 . On est dans les deux cas précédents : Une "bonne" représentation des variables est fournie par les Z_k ,une "bonne" représentation des individus est fournie par les $\mathcal{Y}_k$.

Cependant il ne paraît pas y avoir de lien entre les deux, l'une résultant de l'étude de $\begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix}$,l'autre de celle de $(X_1 \dots X_K)$.

Remarque : Si $O_k = X_k$,c'est-à-dire si on considère comme élément caractéristique de X_k le tableau lui-même,et si on se ramène au schéma de l'Interstructure,



- La diagonalisation de $\mathcal{C}$ donne la représentation des K "variables" dans la base de tous les compromis (qui ici ont tous un sens); mais on peut aussi représenter les "individus",c'est à dire à la fois les variables et les individus de tous les tableaux,à l'aide des vecteurs propres de $\mathcal{X}'\mathcal{X}\varphi$ .On a ainsi un sorte de représentation "moyenne" des variables et des individus tenant compte de plusieurscompromis,et pas seulement du premier .

-On peut aussi effectuer l'Analyse en Composantes Principales du premier compromis(première composante principale de $\mathcal{X}$),puis projeter les variables et les individus de tous les tableaux en éléments supplémentaires.

-Dans le cas particulier de la situation 7 (tableaux de Burt) Une représentation "moyenne" des modalités des deux variables qualitatives est fournie par l'Analyse des Correspondances du compromis : $B = \sum_{k=1}^{K} \alpha_k B_k$. En projetant,de même que précédemment,en éléments supplémentaires,les lois conditionnelles de chaque tableau de contingence,on a une représentation des modalités telles qu'elles sont vues par toutes les études .

CONCLUSION :

L'approche "Interstructure-Compromis-Intrastructure" permet d'étudier des situations très diverses,de façon relativement homogène. Les différences entre les différentes "variantes" de cette approche tiennent essentiellement à la nature des données auxquelles on veut l'appliquer.

Excepté pour la situation 5 (mêmes variables,mêmes individus) le choix du type d'Interstructure,du type de Compromis,et du type d'Intrastructure est imposé par la nature des données.
Pour la situation 5 on peut choisir entre 4 options :

-comparaison des K visions des individus
-comparaison des espaces $\mathcal{E}_k$
-comparaison des K visions des variables
-comparaison des tableaux eux-mêmes .

Ces options donnent des résultats tous à fait différents,car elles mettent en évidence des aspects différents des données .

*
* *

IV : LES MODELES INDSCAL ET IDIOSCAL :

Dans ce chapitre et dans le suivant nous allons présenter des méthodes utilisées surtout dans les pays anglo-saxons pour l'étude conjointe de plusieurs matrices de données.

Les méthodes envisagées ici consistent à postuler un modèle, puis à estimer les paramètres de ce modèle à partir des données. Nous allons exposer rapidement les grandes lignes de cette approche, puis montrer comment l'approche décrite au chapitre III peut rejoindre celle-ci.

1 / LES MODELES :

Ces modèles ont été proposés principalement par CARROLL et CHANG (6), ainsi que TUCKER (30), pour comparer les jugements fournis par K juges sur les mêmes n objets, c'est-à-dire pour comparer K matrices de similarités ou de dissimilarités entre n objets.
Soit D la matrice diagonale des poids des n objets.
On est dans la situation 2) du chapitre II, seule étudiée ici ; à partir des données, on a K matrices W_k, symétriques, semi définies positives, centrées pour D , de dimensions nxn .

Le modèle INDSCAL suppose qu'il existe une matrice X, pxn, et des matrices diagonales positives D_1, D_2, ..., D_K, telles que, pour tout k, $W_k = X'D_kX$.

Le modèle IDIOSCAL suppose qu'il existe une matrice X, pxn, et des matrices symétriques (semi)-définies positives M_1, M_2, ..., M_K, telles que, pour tout k, $W_k = X'M_kX$.

Pour certains auteurs, les M_k ou les D_k doivent être seulement semi-définies positives, alors que pour d'autres, elles doivent être définies positives, et définissent donc véritablement des métriques euclidiennes sur $\mathbb{R}^p$.

2 / SOLUTIONS DE CARROLL-CHANG et TUCKER

2.1 / Pour INDSCAL

Pour estimer X et les D_k, J.D. CARROLL et J.J. CHANG (6) utilisent l'algorithme suivant :

- On choisit $X_G^{(0)}$ et $X_D^{(0)}$, matrices pxn arbitraires, et on estime, par la méthode des moindres carrés, $D_k^{(1)}$, diagonale positive, telle que $X_G^{(0)'} D_k^{(1)} X_D^{(0)}$ soit proche de W_k, pour tout k .

- Puis, conservant $X_G^{(0)}$ et les $D_k^{(1)}$, on estime $X_D^{(1)}$.

- Puis, conservant $X_D^{(1)}$ et les $D_k^{(1)}$, on estime $X_G^{(1)}$, toujours par la méthode des moindres carrés, et ainsi de suite : on a ainsi une suite d'approximations de plus en plus adéquates de W_k.

Soit X_G, D_k, X_D les solutions finales, et $\hat{W}_k = X_G' D_k X_D$.

La symétrie du problème implique qu'il existe une matrice diagonale Δ, telle que $X_G = \Delta X_D$, $\hat{W}_k = X_D' \Delta D_k X_D$: on peut donc prendre finalement comme résultat :

$$\hat{X} = X_D,$$
$$\hat{D}_k = \Delta D_k.$$

L'inconvénient de cette méthode est qu'on ne peut pas être sûr d'arriver à un optimum global. De plus, selon DE LEEUW et PRUZANSKY(10), cet algorithme est en général très lent à converger.

2.2 / Pour IDIOSCAL

L.R. TUCKER (30) propose la méthode suivante : Pour tout k,

Soit Y_k une matrice $p_k$xn de vecteur propres D-orthogonaux de $W_k D$, de norme égale à la racine carrée de la valeur propre correspondante :

$W_k = Y_k' Y_k$. (p_k est le rang de W_k)

Soit $\Lambda_k = Y_k D Y'_k$

Soit C la matrice : $\begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix}$, qu'on peut décomposer, selon le théorème 1 du chapitre I, sous la forme : $C = UY$, où $U = \begin{pmatrix} U_1 \\ \vdots \\ U_k \end{pmatrix}$ est une matrice de dimensions $(\sum_{k=1}^{k} p_k) \times r$, de rang r, et Y est une matrice rxn, de rang r, où r est le rang de C.

On a donc pour tout k : $Y_k = U_k Y$.

$W_k = Y'_k Y_k = Y'(U'_k U_k)\, Y$, où $U'_k U_k$ est symétrique, semi-définie positive.

Si on n'impose pas aux matrices M_k d'être non singulières, on a ainsi une solution exacte pour IDIOSCAL : $X = Y$, $M_k = U'_k U_k$

J. DE LEEUW et S. PRUZANSKY (10) proposent, pour résoudre le problème INDSCAL, un algorithme qui prend comme point de départ une solution de IDIOSCAL, obtenue comme précédemment. Par une généralisation de la méthode de JACOBI pour diagonaliser une matrice symétrique, ils cherchent une matrice T orthonormale telle que la somme des carrés des éléments diagonaux de tous les $T'M_kT$ soit maximale : chaque $T'M_kT$ est le plus proche possible d'une matrice diagonale.
On a alors comme estimation de D_k : $\hat{D}_k = \text{diag}(T'M_kT)$, et comme estimation de X : $\hat{X} = T'Y$.
(Si $T'M_kT = \hat{D}_k$, diagonale, on a $W_k = Y'M_kY = Y'T(T'M_kT)TY$
$= \hat{X}'\hat{D}_k\hat{X}$.)

Cet algorithme, appelé SUMSCAL, converge plus rapidement que celui proposé par CARROLL et CHANG. Sur les exemples traités par DE LEEUW et PRUZANSKY, les deux algorithmes convergent vers des valeurs de X et des D_k très semblables.

2.3 / Représentations graphiques

. Pour les deux modèles, X représente un compromis entre les k études, et fournit donc une représentation "moyenne" des objets.

. Pour représenter chacune des K matrices de similarités ou de dissimilarités, CARROLL et CHANG utilisent les éléments diagonaux des D_k de INDSCAL, et représentent ainsi les K juges par des vecteurs de $\mathbb{R}^p$.

. Pour représenter les n objets "tels qu'ils sont vus par le $k^{\text{è}}$ "juge", on utilise la matrice $(D_k)^{1/2}X$, puisque $W_k = (X'D_k^{1/2})\,(D_k^{1/2}X)$.

3 / SOLUTION ALGEBRIQUE POUR IDIOSCAL ET VALIDITE DU MODELE

Une partie des résultats de ce paragraphe et du suivant se trouvent dans (12, chapitre XIV)et(25), cependant nous y avons apporté des précisions et rectifications.

3,1 / Solution pour IDIOSCAL

On a vu en 2.2) la solution proposée par TUCKER :

Soit $W_k = Y_k'Y_k$, $Y_kDY_k' = \Lambda_k$

$$C = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix},\quad C'C = \sum_{k=1}^{K} W_k = Y'Y, \text{ avec } YDY' = \Lambda$$

On a $W_k = Y'(U_k'U_k)Y$, avec $U_k = Y_kDY'\Lambda^{-1}$

On notera par la suite $Q_k = U_k'U_k = \Lambda^{-1}YDW_kDY'\Lambda^{-1}$: $W_k = Y'Q_kY$,

où Q_k est symétrique semi-définie positive.

Toutes les autres solutions de $W_k = X'M_kX$, où X est de dimension rxn (r = rang de W) sont de la forme $X = AY$, où A est une matrice inversible quelconque, et $M_k = (A')^{-1} Q_k (A)^{-1}$.

Remarque : $\sum_{k=1}^{K} Q_k = \Lambda^{-1} YDWDY' \Lambda^{-1} = \Lambda^{-1} YDY' YDY' \Lambda^{-1} = I$

3.2 / Validité de IDIOSCAL

Jusqu'à présent, on a supposé que le modèle pouvait s'appliquer aux données, et qu'il suffisait d'estimer les paramètres du modèle. Cependant, il peut être gênant de postuler ainsi un modèle a priori, et on peut se demander dans quelle mesure un modèle est applicable aux données, et s'il existe une solution exacte au problème.

* Si on ne cherche que des M_k semi-définies positives, on a vu qu'on peut toujours trouver X et des M_k telles que $W_k = X'M_kX$.

* Cependant, pour que les M_k soient véritablement des métriques euclidiennes, il faut qu'elles soient non singulières. Le problème est alors de savoir si on peut toujours trouver X et des M_k définies positives telles que $W_k = X'M_kX$.

Proposition : Il existe une matrice X et des métriques M_k, telles que $W_k = X'M_k X$ pour tout k, si et seulement si les lignes des Y_k engendrent le même sous-espace vectoriel de $\mathbb{R}^n$.

Démonstration :

* Si $W_k = X'M_kX = Y_k'Y_k$, les lignes de chaque Y_k engendrent le même sous-espace vectoriel de $\mathbb{R}^n$ que les lignes de X.

* Soit $C = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix} = Y'Y$: les lignes de Y engendrent le même sous espace vectoriel de $\mathbb{R}^n$ que la réunion de toutes les lignes de tous les Y_k. Donc si les lignes de Y_k engendrent le même sous-espace vectoriel de $\mathbb{R}^n$, c'est aussi celui engendré par les lignes de Y : rang Y = rang Y_k pour tout k

$= r$

On a alors $W_k = Y'Q_kY$, et Q_k est définie positive (sinon W_k serait de rang < r).

Remarques :

1) Cette proposition signifie qu'on ne peut trouver des solutions M_k définies positives que si les études (W_k, D) sont équivalentes au sens 2) :

pour tout k, $A_k = Y_k'\Lambda_k^{-1}Y_kD = Y'\Lambda^{-1}YD$.

2) Pour vérifier si les projecteurs $A_k = Y_k'\Lambda_k^{-1}Y_kD$ sont égaux, P.H. SCHÖNEMANN (25) propose de calculer $Y'\Lambda^{-1}YDY_k'$ et de le comparer à Y_k'.

Cependant, quelsque soient les tableaux W_k, on a toujours $Y'\Lambda^{-1}YDY_k' = Y_k'$, puisque les lignes de Y engendrent le même sous-espace vectoriel de $\mathbb{R}^n$ que la réunion des lignes de tous les Y_k.

Les différences constatées par SCHÖNEMANN sur un exemple, entre Y_k' et $Y'\Lambda^{-1}YDY_k'$, ne correspondent qu'aux incertitudes sur le calcul de Y, c'est-à-dire sur le calcul des valeurs propres et vecteurs propres de W.
Par contre, les A_k sont égaux si et seulement si $A_kY' = Y'$ pour tout k.

3) Les A_k sont égaux si et seulement si W a le même rang que chacun des W_k. (Sinon le rang de W est supérieur).

3.3 / Liens avec l'approche "Interstructure-Compromis-Intrastructure"

- Pour tester l'égalité des A_k, on peut réaliser l'"Interstructure" des K études, sur la base des produits scalaires $Tr(A_k A_\ell)$. On peut ainsi visualiser les proximités entre les A_k, et éventuellement voir si le modèle IDIOSCAL, avec des M_k définies positives, peut s'appliquer à un certain nombre des W_k.

- On a vu qu'on pouvait construire des solutions (Y, Q_k) à partir de $W = \sum_{k=1}^{K} W_k$. Cependant, tout ce qui a été fait aux paragraphes précédents est conservé si $W = \sum_{k=1}^{K} \alpha_k W_k = Y'Y$, où $\alpha_1, \ldots, \alpha_K$ sont des nombres positifs non nuls quelconques. Le "compromis" Y peut donc résulter de la diagonalisation de n'importe quel compromis (au sens du chapitre III) entre les W_k, à condition qu'aucun α_k ne soit nul. Si il y a des α_k nuls, on ne peut reconstituer W_k sous la forme $W_k = Y'Q_kY$ que pour les indices k tels que α_k est non nul.

- On a $W_k = Y'Q_kY = Y'\Lambda^{-1}YDY'_kY_kDY'\Lambda^{-1}Y$.

Si pour représenter la vision "moyenne" des n objets, on se limite aux premières lignes de Y (notées Y_o) on a alors une reconstitution approchée des W_k sous la forme :

$\hat{W}_k = (Y'_o\Lambda_o^{-1}Y_oDY'_k)\,(Y_kDY'_o\Lambda_o^{-1}Y_o) = \hat{Y}'_k\hat{Y}_k$, où $\hat{Y}_k$ est la représentation des individus dans l'Intrastructure, selon la 1ère méthode décrite en III-4.1.

- Si $W_k = Y'Q_kY$ pour tout k,

$Tr\,(W_kDW_\ell D) = Tr\,(Y'Q_KYDY'Q_\ell YD) = Tr\,(Q_k\Lambda Q_\ell\Lambda) = Tr\,(M_kM_\ell)$, en posant, pour tout k, $M_k = \Lambda^{1/2}\,Q_k\,\Lambda^{1/2}$: réaliser"l'Interstructure" des K études, sur la base des produits scalaires $Tr\,(W_kDW_\ell D)$, revient à comparer les matrices M_k.

Comme $W_k = (Y'\Lambda^{-1/2})\ M_k\ (\Lambda^{-1/2}Y)$, cela revient à considérer qu'il y a, pour les k jugements, une partie commune donnée par $\Lambda^{-1/2}Y$ (vecteurs propres normés de WD), et que M_k représente la façon dont le kième juge a noté les n objets, et est caractéristique de sa "vision" des n objets.

4 / SOLUTION ALGEBRIQUE POUR INDSCAL ET VALIDITE DU MODELE

4.1 / Solution

On suppose qu'il existe X, matrice pxn de rang r, avec $r \leqslant p \leqslant n$, et des matrices diagonales définies positives $D_1, D_2, \ldots, D_K$, telles que $W_k = X'D_kX$ pour tout k.

On a d'autre part, comme au paragraphe précédent : $W_k = Y'Q_kY$, où Y est une matrice rxn, de rang r :

$W_k = Y'Q_kY = X'D_kX$. Donc il existe H, matrice pxr, telle que pour tout k, $H'D_kH = Q_k$, et $X = HY$.

Comme $\sum_{k=1}^{K} Q_k = I$ par construction (voir remarque de 3.1), en posant

$\Delta = \sum_{k=1}^{K} D_k$, on a $H'\Delta H = I$.

On a donc $X = \Delta^{-1/2}(\Delta^{1/2}H)Y$, avec $(H'\Delta^{1/2})\ (\Delta^{1/2}H) = I$.

Si on se limite aux X de dimensions rxn (r = rang de W = rang de chaque W_k), toutes les solutions sont alors de la forme : $W_k = X'D_kX$, avec $X = \Delta^{-1/2}TY$, où Δ est une matrice diagonale définie positive quelconque, et T est une matrice orthonormale telle que $Q_k = T'\ (\Delta^{1/2}D_k\Delta^{1/2})\ T$; T est

donc formée de vecteurs propres orthonormés de Q_k.

Remarque :

On pourrait être tenté, pour résoudre INDSCAL, d'estimer les matrices $\hat{D}_k$ diagonales, telles que $Y'\hat{D}_kY$ soit proche de W_k, pour tout k, par la méthode des moindres carrés :

$$\text{or } ||W_k - Y'\hat{D}_kY||^2_{\varphi_{DD}} = ||Y'(Q_k-\hat{D}_k)Y||^2_{\varphi_{DD}} = \text{Tr } (Y'(Q_k-\hat{D}_k)\, YDY'\, (Q_k-\hat{D}_k)\, YDY')$$

$$= \text{Tr } ((Q_k-D_k)\Lambda)^2 = \sum_{i=1}^{r} ((Q_k)_{ii}-(\hat{D}_k)_{ii})^2\lambda_i^2 + \sum_{i\neq j} (Q_k)_{ij}^2\lambda_i^2$$

le minimum est atteint pour $(\hat{D}_k)_{ii} = (Q_k)_{ii}$: $\hat{D}_k$ a pour éléments diagonaux les éléments diagonaux de Q_k.

Cependant, même si le modèle INDSCAL est exact, on peut avoir ainsi des résultats très différents de la solution exacte :

la matrice $Q_k = T'\,(\Delta^{-1/2}D_k\Delta^{-1/2})\,T$ peut être très différente d'une matrice diagonale.

4.2 / Validité du modèle

Si on ne cherche que des D_k semi-définies positives, on peut toujours trouver X et des D_k telles que $W_k = X'D_kX$, en prenant par exemple

$$X = \begin{pmatrix} Y_1 \\ \vdots \\ Y_K \end{pmatrix}, \text{ et } D_k = \begin{pmatrix} 0 & & \\ & I_{p_k} & 0 \\ 0 & & \ddots \\ & & & 0 \end{pmatrix}$$

Si on se limite aux D_k définies positives, le problème est de savoir si on peut toujours trouver X et des D_k telles que $W_k = X'D_kX$.

Proposition : Il existe une matrice X et des métriques diagonales D_k, telles que $W_k = X'D_kX$ pour tout k, si et seulement si les lignes de Y_k engendrent le même sous-espace vectoriel de $\mathbb{R}^n$, et si les matrices Q_k ont un même système orthonormal de vecteurs propres (c'est-à-dire si elles ont les mêmes sous-espaces vectoriels propres).

Démonstration :

* S'il existe X et des D_k diagonales définies positives telles que $W_k = X'D_kX$, les A_k sont égaux, et valent $A_Y = Y'\Delta^{-1}YD$ (on a un cas particulier de IDIOSCAL). Les W_k sont donc tous de même rang, celui de Y : r. On peut alors, sans perte de généralité, considérer que X est de dimensions rxn, de rang r. En construisant en 4.1 les solutions du problème, on a vu qu'il existe alors une matrice T orthonormale telle que :

$Q_k = T'(\Delta^{-1/2}D_k\Delta^{1/2})\ T$, les Q_k ont tous un même système orthonormal de vecteurs propres, fourni par T.

* Inversement, si les A_k sont égaux, $W_k = Y'Q_kY$, et les Q_k sont définis positifs. Si de plus il existe H, matrice orthonormale de vecteurs propres de chaque Q_k, on a $Q_k = H'D_kH$, où D_k est diagonale définie positive :

$W_k = (Y'H')\ D_k\ (HY)$.

Application :

Pour tester la validité du modèle, on peut d'abord comparer les A_k. Si les A_k sont égaux, il reste à trouver T orthonormale telle que $T'Q_kT$ soit

diagonale pour tout k. Il ne semble pas y avoir de moyen simple de connaître l'existence ou la non-existence d'une telle matrice T. On a seulement le résultat suivant :

Proposition : Si les Q_k n'ont que des valeurs propres simples, il existe un système orthonormal de vecteurs propres pour tous les Q_k si et seulement si pour tout k et tout ℓ, $Q_k Q_\ell = Q_\ell Q_k$.

Démonstration :

* Si $Q_k = H' D_k H$, où H est orthonormale, $Q_k Q_\ell = H' D_k D_\ell H = Q_\ell Q_k$

* Si $Q_k Q_\ell = Q_\ell Q_k$, soit u un vecteur propre de Q_k.

$Q_k u = \lambda u \Rightarrow Q_\ell Q_k u = Q_k(Q_\ell u) = \lambda(Q_\ell u)$: $Q_\ell u$ est vecteur propre de Q_k pour λ, donc il existe un nombre μ tel que $Q_\ell u = \mu u$: u est aussi vecteur propre de Q_ℓ . De façon identique, on voit que si u est vecteur propre de Q_ℓ, il est vecteur propre de Q_k. Donc Q_k et Q_ℓ ont les mêmes vecteurs propres .

Remarque :

Si on a n'a que deux matrices W_1 et W_2, on a $W_1 = Y' Q_1 Y$, $W_2 = Y' Q_2 Y$, avec $Q_1 + Q_2 = I$. Soit H un système orthonormal de vecteurs propres de Q_1 : $Q_1 = H D_1 H'$; $Q_2 = H(I - D_1)H'$.

Donc dans le cas de deux "juges", INDSCAL a une solution si et seulement si IDIOSCAL en a une, c'est-à-dire si et seulement si les deux études sont équivalentes au sens 2) .

4.3 / Liens avec l'approche "Interstructure-Compromis-Intrastructure"

Comme INDSCAL est un cas particulier d'IDIOSCAL, tout ce qui a été dit en 3.3 est encore valable.

* La comparaison des A_k par l'Interstructure est utile, mais non suffisante, pour connaître la validité du modèle.

* pour le compromis : mêmes remarques qu'en 3.3.

* si $W_k = X'D_kX$, $Tr(W_k D\ W_\ell D) = Tr(D_k XDX' D_\ell XDX')$.

CARROLL et CHANG réalisent une sorte d'"Interstructure" des K juges en les représentant par les vecteurs de $\mathbb{R}^n$ obtenus en prenant les éléments diagonaux des D_k, et justifient cette représentation par le fait que si XDX' est près de l'identité, $Tr(W_k DW_\ell D)$ est proche de TrD_kD_ℓ. Cependant, on a vu en 4.1 que X est de la forme $X = \Delta^{-1/2}TY$, avec $Q_k = T'(\Delta^{1/2}D_k\Delta^{-1/2})T$: $XDX' = \Delta^{-1/2}T\Lambda T'\Delta^{-1/2}$, a priori rien ne permet d'assimiler XDX' à une matrice identité, ni même à une matrice diagonale, à moins que les Q_k soient proches de matrices diagonales.

CONCLUSION

La méthode IDIOSCAL, bien que formulée au départ de façon très différente (modèle postulé a priori, et paramètres estimés ensuite) de l'approche "Interstructure-Compromis-Intrastructure", rejoint celle-ci sur bien des points.

La méthode INDSCAL, bien que comportant aussi les 3 étapes : représentation globale (par les éléments diagonaux de D_k), compromis (X), représentation détaillée (par $D_k^{1/2}X$), conduit à des résultats sensiblement différents.

Cela vient du fait que ce modèle n'est applicable,même approximativement, que si les données ont une structure très particulière.Il vaut mieux dans ce cas rechercher une méthode algorithmique de résolution,comme celle proposée par J.DE LEEUW et S.PRUZANSKY.

L'intérêt de la méthode INDSCAL est que, si le modèle s'applique approximativement, les résultats sont facilement interprétables. Les K jugements sont basés sur la même matrice X, les différences proviennent uniquement de la différence de pondération des axes.

*

* *

V : LES METHODES PROCRUSTEENNES :

Lorsqu'on dispose de plusieurs matrices de données résultant de l'observation des mêmes individus, on peut comparer les liens entre les individus, d'une étude à l'autre, en comparant les "$W_K D$" associés, comme on l'a vu au chapitre III.

Une autre approche, appelée dans la littérature américaine "Procrustes Analysis" consiste à rechercher comment transformer les matrices initiales pour les rapprocher, tout en conservant les positions relatives des individus pour une même étude.

Nous allons voir d'abord comment on peut procéder dans le cas de deux tableaux de données, puis pour K tableaux, et comparer cette approche aux précédentes.

1 / CAS DE DEUX MATRICES DE DONNEES

1.1 / Le problème et sa solution

Soit deux triplets (X_1, M_1, D) et (X_2, M_2, D), où X_1 et X_2 sont de dimensions $p_1 x n$ et $p_2 x n$, avec $p_1 \geq p_2$, X_1 et X_2 centrées pour D. M_1 et M_2 sont deux métriques euclidiennes sur $\mathbb{R}^{p1}$ et $\mathbb{R}^{p2}$.

Problème : On cherche H réalisant le minimum de $\|X_1 - HX_2\|^2_{M_1,D}$, parmi les matrices H, $p_1 x p_2$, telles que $H'M_1H = M_2$.

Cela signifie qu'on veut remplacer (X_2, M_2, D) par (HX_2, M_1, D), étude équivalente au sens 1 (même W), et qui soit le plus proche possible de X_1 pour la distance induite par φ_{M_1D}, c'est-à-dire telle que la somme pondérée (par D) des carrés des distances entre points homologues (individus de $\mathbb{R}^{p_1}$) soit minimale .

Solution : Ce problème est résolu dans (26), à l'aide de la technique des multiplicateurs de Lagrange, pour le cas particulier où $p_1 = p_2$.

et $M_1 = M_2 = I$.

P. BOURGEOIS (2) le résout d'une façon plus immédiate, mais toujours dans le cas où $p_1 = p_2$. On peut cependant le résoudre directement dans le cadre général de l'énoncé ci-dessus :

$$||X_1 - HX_2||^2_{\varphi_{M_1 D}} = ||X_1||^2_{\varphi_{M_1 D}} + ||HX_2||^2_{\varphi_{M_1 D}} - 2\langle X_1, HX_2\rangle_{\varphi_{M_1 D}} = \mathrm{Tr}W_1 D + \mathrm{Tr}W_2 D - 2\mathrm{Tr}(X_1 D X_2' H' M_1$$

Le problème revient à chercher H, tel que $H'M_1H = M_2$, maximisant : $\mathrm{Tr}(X_1 D X_2' H' M_1)$.

$X_1 D X_2'$ est une matrice $p_1 \times p_2$, que l'on peut décomposer, selon le théorème 1 du chapitre I, sous la forme :

$X_1 D X_2' = U\Delta V'$, avec $U'M_1U = I$, $V'M_2V = I$, $\Delta = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \end{pmatrix}$, $\lambda_i > 0$ pour

$i = 1, \ldots r$, $U = (U_1 \ldots U_r)$ $V = (V_1 \ldots V_r)$, où r est le rang de $X_1 D X_2'$.

Alors $\mathrm{Tr}(X_1 D X_2' H' M_1) = \mathrm{Tr}(U\Delta V' H' M_1) = \mathrm{Tr}(\Delta V' H' M_1 U) = \sum_{i=1}^{r} \lambda_i \langle HV_i, U_i\rangle_{M_1}$

$$\leq \sum_{i=1}^{r} \lambda_i ||HV_i||_{M_1} ||U_i||_{M_1} = \sum_{i=1}^{r} \lambda_i ||V_i||_{M_2} ||U_i||_{M_1} = \sum_{i=1}^{r} \lambda_i$$

Le maximum est atteint pour tout H tel que $HV = U$, et vaut $\sum_{i=1}^{r} \lambda_i = \mathrm{Tr}\Delta$.

* Si $X_1 D X_2'$ est de rang p_2, H est défini de façon unique par $HV = U$.

(U et V étant fixés).

$(V'M_2)V = I \Rightarrow V(V'M_2) = I \Rightarrow$ $H = UV'M_2$, et on a bien :

$H'M_1H = M_2VU'M_1UV'M_2 = M_2$.

* Si X_1DX_2' est de rang inférieur à p_2, H n'est pas unique :

Soit $\tilde{V} = (V\ V_{r+1}\ldots V_{p_2})$ telle que $\tilde{V}'M_2\tilde{V} = I$. Toute solution H est définie de façon unique par ses valeurs prises sur la base $\{V_1,\ldots,V_r,V_{r+1},\ldots V_{p_2}\}$:

$H\tilde{V} = (HV\ \ HV_{r+1}\ldots HV_{p_2}) = (U\ \ U_o)$, et H doit vérifier : $H'M_1H = M_2$,

ce qui est équivalent à : $\tilde{V}'H'M_1H\tilde{V} = \tilde{V}'M_2\tilde{V}$:

$= I$, ou $\begin{pmatrix} U'M_1U & U'M_1U_o \\ U_o'M_1U & U_o'M_1U_o \end{pmatrix} = I$: H est

solution si et seulement si $\tilde{U} = (U\ U_o)$ est M_1-orthonormale. Alors $H\tilde{V} = \tilde{U}$,

$H = \tilde{U}\tilde{V}'M_2$

Donc toutes les solutions sont de la forme $H = \tilde{U}\tilde{V}'M_2$, où $\tilde{U}$ et $\tilde{V}$ sont deux matrices M_1 et M_2-orthonormales obtenues en complétant U et V, de façon quelconque.

Remarque : Si X_1DX_2' est de rang p_2, $H = UV'M_2 = U\Delta V'M_2$ $(V\Delta^{-1}V'M_2)$,

et $(V\Delta^{-1}V'M_2)^2 = V\Delta^{-2}V'M_2 = (V\Delta^2V'M_2)^{-1}$:

$$H = X_1DX_2'M_2(X_2DX_1'M_1X_1DX_2'M_2)^{-1/2}$$

et $||X_1 - HX_2||^2_{\varphi_{M_1D}} = \mathrm{Tr}W_1D + \mathrm{Tr}W_2D - 2\mathrm{Tr}((X_2DX_1'M_1X_1DX_2'M_2)^{1/2})$

Proposition : Une condition nécessaire et suffisante d'optimum est que $X_1DX_2'H'$ soit symétrique et semi-définie positive.

En effet :

* à l'optimum $X_1DX_2'H' = (\tilde{U}\begin{pmatrix}\Delta & o\\ o & o\end{pmatrix}\tilde{V}')(\tilde{U}\tilde{V}'M_2)' = \tilde{U}\begin{pmatrix}\Delta & o\\ o & o\end{pmatrix}\tilde{U}' = U\Delta U'$, symétrique et semi-définie positive.

* Si $X_1DX_2'H'$ est symétrique et semi-définie positive, les valeurs propres de $X_1DX_2'H'M_1$ sont positives, soit $\hat{\Delta}$ la matrice diagonale de ces valeurs propres : $Tr(X_1DX_2'H'M_1) = Tr\hat{\Delta}$.

$\hat{\Delta}^2$ est la matrice diagonale des valeurs propres de $(X_1DX_2'H'M_1)^2 = X_1DX_2'H'M_1HX_2DX_1'M_1$, (puisque $X_1DX_2'H'$ est symétrique)

$(X_1DX_2'H'M_1)^2 = (X_1DX_2')\,M_2\,(X_2DX_1')\,M_1$, dont les valeurs propres sont les éléments diagonaux de Δ^2 ; $Tr(\hat{\Delta}) = Tr\Delta$: on est à l'optimum.

Remarque :

Lorsque p_1 est supérieur à p_2, GOWER (14), ainsi que SCHÖNEMANN et CARROLL (26), se ramènent au cas $p_1 = p_2$, en remplaçant X_2 par $\hat{X}_2 = \begin{pmatrix}X_2\\ o\end{pmatrix}$.

On se limite ici au cas $M_1 = I$ et $M_2 = I$.

$X_1DX_2' = U\Delta V'$, avec $U'U = I$, $V'V = I$. $X_1D\hat{X}_2' = U\Delta(V'\ 0)$, avec $(V'o)\begin{pmatrix}V\\ o\end{pmatrix} = I$, H est défini par $H\begin{pmatrix}V\\ o\end{pmatrix} = U$: H peut être obtenue en complétant de façon quelconque H_o tel que $H_oV = U$, solution du problème précédent, et on a $H\hat{X}_2 = H_oX_2$. Cette méthode, consistant à "ajouter des zéros", conduit au résultat du problème précédent, et est donc beaucoup moins arbitraire qu'il ne le paraît au premier abord.

1.2 / Problème "Oblique Procrustes"

Pour traiter le cas où p_1 est supérieur à p_2, on cherche à résoudre ici un problème différent : On cherche une matrice T, telle que diag (TT') = I, et $||X_1-TX_2||^2_{\varphi_{ID}}$ soit minimale : cela revient à chercher successivement p_1 combi-

naisons linéaires des p_2 variables de X_2. Bien que nommé "Oblique Procrustes Problem", ce problème est fondamentalement différent de celui de "Orthogonal Procrustes Problem" résolu précédemment.

* La solution exacte de ce problème, assez complexe, est donnée par TEN BERGE et NEVELS (28).

* La pratique généralement employée est la suivante :

Pour chercher t_i tel que $||x_i - t_i X_2||_D^2$ soit minimum, avec $t_i' D t_i = 1$, on "enlève" la contrainte $t_i' t_i = 1$, on trouve $t_i = x_i D X_2' (X_2 D X_2')^{-1}$, et on normalise les t_i obtenus.

Sans cette normalisation, on a : $T = (X_2 D X_2')^{-1} X_2 D X_1'$: la recherche de T revient à effectuer l'analyse en composantes principales de X_2 par rapport à X_1.

(Voir (24) et (12) par exemple).

1.3 / Extension de la méthode

Problème : (X_1, M_1, D) et (X_2, M_2, D) étant définis comme précédemment, on cherche une matrice H, $p_1 x p_2$, telle que $H' M_1 H = M_2$, un scalaire $c > 0$, et un vecteur t de $\mathbb{R}^p$, tels que :

$||X_1 - (cHX_2 + t\underset{\sim}{1}')||^2_{\varphi_{M_1} D}$ soit minimal.

Cela signifie qu'on veut remplacer X_2 par $cHX_2 + t\underset{\sim}{1}'$, translaté de cHX_2, équivalent à X_2 au sens 1' (W proportionnels).

Solution :

* Etant données les propriétés du centre de gravité, $||X_1 - cHX_2 - t\underset{\sim}{1}'||^2$ est minimum lorsque les centres de gravité de X_1 et de cHX_2 coïncident, ce qui est le cas ici puisque X_1 et X_2 sont centrés, donc $t = 0$.

$$* \; ||X_1 - cHX_2||^2_{\varphi_{M_1}D} = c^2 \, TrW_2D - 2c \, \langle X_1, HX_2 \rangle + TrW_1D.$$

Le minimum est atteint pour $c = \dfrac{\langle X_1, HX_2 \rangle}{TrW_2D}$, et vaut $TrW_1D - \dfrac{\langle X_1, HX_2 \rangle^2}{TrW_2D}$.

H doit rendre maximum $\langle X_1, HX_2 \rangle_{\varphi_{M_1}D}$: on a les mêmes valeurs de H que pour le problème exposé en 1.1.

Alors $c = \dfrac{Tr\Delta}{TrW_2D}$, et $||X_1 - cHX_2||^2_{\varphi_{M_1}D} = TrW_1D - \dfrac{(Tr\Delta)^2}{TrW_2D}$.

Remarque :

Quand $p_1 = p_2$, et $M_1 = M_2 = M$,

Si on veut approcher de la même manière X_1 de X_2, au lieu d'approcher X_2 de X_1, on trouve H : $V'U \; \tilde{V}'\tilde{U}M = M\,H'M = (H)^{-1}$

$$c = \frac{Tr\Delta}{TrW_1D}.$$

2 / MESURE DE L'AJUSTEMENT - "INTERSTRUCTURE"

On a vu en 1.1 que : $\min_{H'M_1H = M_2} ||X_1 - HX_2||^2_{\varphi_{M_1}D} = Tr(W_1D) + Tr(W_2D) - 2tr\Delta$,

où Δ^2 est la matrice diagonale dont les éléments diagonaux sont les valeurs propres de $X_1DX_2'M_2X_2DX_1'M_1$, c'est-à-dire les valeurs propres de W_1DW_2D.

Si on dispose (situation 1 du chapitre II) de K études (X_k, M_k, D) portant sur les mêmes individus, où X_k est une matrice $p_k$xn centrée pour D,

on peut définir, pour tout couple (k,ℓ),

$dis^2(W_k,W_\ell) = Tr(W_kD) + Tr(W_\ell D) - 2Tr(\Delta_{k\ell})$, où $\Delta^2_{k\ell}$ est la matrice diagonale dont les éléments diagonaux sont les valeurs propres de $W_kD\,W_\ell D$.

On montre facilement ((2), ou (14)) que dis définit une distance sur l'ensemble des W_k, mais que cette distance n'est pas euclidienne.

D'autre part on a vu en 1.3 que $\min_{\substack{H'M_1H = M_2 \\ c > 0}} \varphi_{M_1D}(X_1 - cHX_2)^2 = TrW_1D - \frac{(Tr\Delta)^2}{TrW_2D}$

cet indice n'étant pas symétrique, on préfère utiliser l'indice de dissimilarité :

$dis^2\left(\frac{W_k}{TrW_kD}, \frac{W_\ell}{TrW_\ell D}\right) = 2\left(1 - \frac{Tr\Delta k_\ell}{\sqrt{TrW_kD\ TrW_\ell D}}\right)$, ou de façon équivalente, l'indice de similarité : $\frac{Tr\Delta k_\ell}{\sqrt{TrW_kD\ TrW_\ell D}}$, noté cor $(W_kD, W_\ell D)$, utilisé par LINGOES et SCHÖNEMANN (21). Cet indice est compris entre 0 et 1. Il est nul si et seulement si $Tr(W_kDW_\ell D) = 0$, c'est-à-dire $X_kDX'_\ell = 0$. Il est égal à 1 si et seulement si $\frac{W_k}{TrW_kD} = \frac{W_\ell}{TrW_\ell D}$, c'est-à-dire si et seulement si (X_k, M_k, D) et (X_ℓ, M_ℓ, D) sont équivalents au sens 1') (W proportionnels).

On peut réaliser l'"interstructure" des K études sur la base des distances $dis^2(W_k, W_\ell)$, ou des similarités cor $(W_kD, W_\ell D)$, en diagonalisant la matrice $\mathcal{C}$, KxK, symétrique, semi-définie positive, construite à partir de la matrice des $dis^2(W_k, W_\ell)$ ou des cor $(W_kD, W_\ell D)$, selon la méthode décrite au chapitre II.1 (situation 2)

Si on décrit des "études-variables" (Voir chapitre III-2.2), on diagonalise une approximation semi-définie positive de la matrice $\mathcal{C}$, KxK, d'élément $\mathcal{C}_{k\ell} = \mathrm{Tr}\Delta_{k\ell}$,

ou bien $\mathcal{C}_{k\ell} = \dfrac{\mathrm{Tr}\Delta k\ell}{\sqrt{\mathrm{Tr}W_kD\ \mathrm{Tr}W_\ell D}} = \mathrm{cor}\ (W_kD,\ W_\ell D)$.

Cette méthode est comparable à l'"Interstructure" de la méthode STATIS, où l'on diagonalise la matrice des $\mathrm{Tr}\ (W_kD\ W_\ell D) = \mathrm{Tr}\ (\Delta_{k\ell})^2$, ou bien la matrice

des $\mathrm{Rv}\ (W_kD,\ W_\ell D) = \dfrac{\mathrm{Tr}(\Delta_{k\ell})^2}{\sqrt{\mathrm{Tr}(W_kD)^2\ \mathrm{Tr}(W_\ell D)^2}}$

On remarque en particulier que : $\mathrm{Rv}(W_kD,W_\ell D) = 0 \iff \mathrm{cor}\ (W_kD,W_\ell D) = 0$

$\mathrm{Rv}(W_kD,W_\ell D) = 1 \iff \mathrm{cor}\ (W_kD,W\ D) = 1$

$||W_k-W_\ell||^2_{\varphi_{DD}} = 0 \iff \mathrm{dis}^2\ (W_k,\ W_\ell) = 0.$

Cependant, la méthode STATIS présente l'avantage de fournir directement une matrice $\mathcal{C}$ semi-définie positive, et une image euclidienne exacte des K études.

3 / ANALYSE PROCUSTEENNE GENERALISEE

3.1 / Le problème et sa solution

On dispose de K triplets (X_k,M,D), où les X_k sont des matrices pxn, centrées pour D, et M une métrique sur $\mathbb{R}^p$. On est dans la situation 5) du chapitre II.

On associe à chaque triplet un poids p_k tel que $\sum_{k=1}^{K} p_k = 1$.

Problème : On cherche $H_1, H_2, \ldots, H_K$ réalisant le minimum de

$$\sum_{k=1}^{K} \sum_{\ell=1}^{K} p_k p_\ell \, ||H_k X_k - H_\ell X_\ell||^2_{\varphi_{MD}}$$

sous les contraintes : $H_k' M H_k = M$ pour $k = 1, \ldots, K$.

Cela signifie qu'on cherche à remplacer chaque triplet par un triplet équivalent au sens 1) (même W), de telle sorte que les tableaux obtenus soient globalement le plus proches possible les uns des autres.

Solution :

* On remarquera d'abord que : $||H_k X_k - H_\ell X_\ell||^2_{\varphi_{MD}} = ||X_k||^2_{\varphi_{MD}} + ||X_\ell||^2_{\varphi_{MD}} - 2 \langle H_k X_k, H_\ell X_\ell \rangle_{\varphi_{MD}}$

$$\sum_{k=1}^{K} \sum_{\ell=1}^{K} p_k p_\ell \, ||H_k X_k - H_\ell X_\ell||^2_{\varphi_{MD}} = 2 \left(\sum_{k=1}^{K} p_k ||X_k||^2_{\varphi_{MD}} - \left|\left| \sum_{k=1}^{K} p_k H_k X_k \right|\right|^2_{\varphi_{MD}} \right)$$

$$= 2 \sum_{k=1}^{K} p_k \, ||H_k X_k - \sum_{\ell=1}^{K} p_\ell H_\ell X_\ell||^2 .$$

Le problème revient ainsi à chercher des H_k tels que $\sum_{k=1}^{K} p_k H_k X_k$ soit d'inertie maximum, ou tels que $\sum_{k=1}^{K} p_k \, ||H_k X_k - \sum_{\ell=1}^{K} p_\ell H_\ell X_\ell||^2$ soit minimum.

$\sum_{\ell=1}^{K} p_k H_k X_k$ joue le rôle d'un "compromis" entre les K études ; c'est le compro-

mis n° 1 (du chapitre III) des tableaux $H_k X_k$, chaque tableau $H_k X_k$ étant considéré comme le plus représentatif de W_k.

* Il ne semble pas y avoir de solution algébrique au problème. Toutes les méthodes de résolution actuelles sont des méthodes algorithmiques.
On a les deux résultats suivants, démontrés par P. BOURGEOIS (2), ainsi que J.M.F. TEN BERGE (28) :

1) Une condition nécessaire d'optimum est que chaque X_k soit à distance minimale, au sens "dis", de la moyenne des autres X_k, c'est-à-dire que chaque $H_k X_k D \left(\sum_{\ell \neq k} p_\ell H_\ell X_\ell \right)'$ soit symétrique et semi-défini positif. Cette condition n'est pas suffisante.

2) La condition nécessaire ci-dessus étant remplie, une condition suffisante d'optimum est que la matrice :

$$\Psi = \begin{pmatrix} S_1 & -V_{12} \cdots -V_{1K} \\ \vdots & \vdots \\ -V_{K1} \cdots \cdots & S_k \end{pmatrix}$$ soit symétrique et semi-définie positive,

où $V_{k\ell} = p_k p_\ell \, (H_k X_k D X'_\ell H'_\ell)$

$S_k = \sum_{k \neq \ell} V_{k\ell}$. Cette condition n'est pas nécessaire.

P. BOURGEOIS et J.M.F. TEN BERGE donnent deux algorithmes itératifs légèrement différents, convergeant vers une solution qui vérifie la condition nécessaire d'optimum, et qui ne correspond pas forcément à l'optimum global. Cependant, selon P.BOURGEOIS, la condition suffisante est vérifiée très fréquemment en pratique, et sert à tester si la procédure itérative converge vers l'optimum global.

3.2 / Extensions de la méthode

* J.C. GOWER (14) cherche simultanément $H_1, H_2, \ldots, H_K$, $T_1, T_2, \ldots, T_K$ et $C_1, C_2, \ldots, C_K$, minimisant :

$$\sum_{k=1}^{K} \sum_{\ell=1}^{K} ||C_k H_k X_k + T_k - (C_\ell H_\ell X_\ell + T_\ell)||^2_{\varphi_{MD}}, \text{ où } H_k' M H_k = M, C_k > 0, \text{ et}$$

$T_k = t_k \underset{\sim}{1}'$, translation, avec $\sum_{k=1}^{K} C_k^2 ||X_k||^2 = \sum_{k=1}^{K} ||X_k||^2$

Solution :

De même qu'en 1.3, l'optimum (pour les T_k) est atteint lorsque les centres de gravité des $C_k H_k X_k$ coïncident, ce qui est le cas ici puisque les X_k sont centrés.

donc $t_1 = t_2 = \ldots = t_k = 0$.

GOWER recherche les H_k, et les C_k, alternativement, par une procédure itérative. La recherche des H_k se fait de la manière ci-dessus, mais, ainsi que le souligne J.M.F. TEN BERGE (28), celle des C_k n'est pas correcte.

Recherche des C_k :

Supposons que les H_k soient connus, et soit $\hat{X}_k = H_k X_k$.

On veut minimiser $\sum_{k=1}^{K} \sum_{\ell=1}^{K} ||C_k \hat{X}_k - C_\ell \hat{X}_\ell|| = 2K \sum_{k=1}^{K} C_k^2 ||\hat{X}_k||^2 - 2 \sum_{k=1}^{K} \sum_{\ell=1}^{K} C_k C_\ell \langle \hat{X}_k, \hat{X}_\ell \rangle$,

sous la contrainte : $\sum_{k=1}^{K} C_k^2 ||X_k||^2 = \sum_{k=1}^{K} ||X_k||^2$.

En posant $d' = (C_1||X_1|| \ldots C_k||X_K||)$, on doit maximiser $d'Rd$, sous la contrainte

$d'd = \sum_{k=1}^{K} ||X_k||^2$, où R est la matrice KxK des $\dfrac{\langle \hat{X}_k, \hat{X}_\ell \rangle}{||X_k||\ ||X_\ell||}$: d est vecteur

propre de R, associé à la plus grande valeur propre, de norme $\sqrt{\sum_{k=1}^{K} ||X_k||^2}$.

On remarque alors que $\dfrac{1}{\sum_{k=1}^{K} ||X_k||^2} \sum_{k=1}^{K} C_k \tilde{X}_k$ est le compromis des $\dfrac{\tilde{X}_k}{||X_k||}$,

au sens du compromis n° 2 défini en III-3.

Donc, selon que l'on recherche uniquement des H_k, ou bien des H_k et des C_k, on obtient à l'optimum, comme compromis, le compromis n° 1 entre les $H_k X_k$, ou bien le compromis n° 2 entre les $\dfrac{H_k X_k}{||X_k||}$.

(ou entre les $H_k X_k$ si on modifie la contrainte imposée aux C_k, en remplaçant

$$\sum_{k=1}^{K} C_k^2 ||X_k||^2 = \sum_{k=1}^{K} ||X_k||^2 \quad \text{par} \quad \sum_{k=1}^{K} C_k^2 = 1).$$

On peut aussi étendre l'analyse procustéenne généralisée au cas où les X_k n'ont pas le même nombre de variables, et où $M_k = I_{p_k}$ pour tout k :

on remplace chaque matrice X_k par $\begin{pmatrix} X_k \\ o \\ \vdots \\ o \end{pmatrix}$, matrice pxn, où $p = \max_{k=1,2,\ldots,K} \{ p_k \}$

3.3 / Représentations détaillées :

Lorsqu'on dispose de $H_1, H_2, \ldots, H_K$ réalisant le minimum de :

$\sum_{k=1}^{K} \sum_{l=1}^{K} p_k p_l ||H_k X_k - H_l X_l||^2_{\varphi_{MD}}$, pour toute matrice H telle que H'MH=M, les $HH_1, HH_2, \ldots, HH_K$ réalisent aussi ce minimum. En particulier, on peut ainsi représenter simultanément les individus de tous les tableaux en rapportant les coordonnées de tous les $H_k X_k$ au repère formé par les axes factoriels du compromis $\sum_{k=1}^{K} p_k H_k X_k$.

De plus les H_kX_k sont à distance $dis(W_k, W)$ du compromis, où WD est l'opérateur associé au compromis : $\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l X_k' H_k' M H_l X_l D$ (condition nécessaire d'optimum).

WD est tel que $\sum_{k=1}^{K} dis^2(W_k, W)$ soit minimum.

On a vu au chapitre III la méthode utilisée dans STATIS pour l'Intrastructure: $\tilde{Y}_k = \Lambda^{-1/2} Y D Y_k' \Lambda_k^{-1/2} Y_k$. Pour tout k,

$\tilde{Y}_k DY'\Lambda^{-1/2} = \Lambda^{-1/2} YDY_k'\Lambda_k^{-1/2} Y_k DY'\Lambda^{-1/2}$ est symétrique et semi définie positive, $\tilde{Y}_k$ est donc à la distance $dis(W_kD, \bar{W}D)$ de $\Lambda^{-1/2}Y$ où $\bar{W}D = (\Lambda^{1/2}Y)'(\bar{\Lambda}^{1/2}Y) = Y'\Lambda^{-1}Y$.

$\tilde{Y}_k$ est donc un tableau, équivalent à Y_k au sens 1) (même W), à distance minimale de $\Lambda^{-1/2}Y$.

CONCLUSION :

Nous avons tenté de mettre en évidence, dans la description des méthodes procustéennes, les liens qui existent entre ces méthodes et l'approche décrite au chapitre III.

On retrouve toujours les trois grandes étapes: représentation globale (des distances $dis(W_kD, W_lD)$), compromis (qui sont en fait les compromis "vus" au chapitre III, mais utilisés après transformation des données initiales), et représentation détaillée de toutes les données, après transformation.

Cette approche a permis aussi d'apporter une justification supplémentaire à l'utilisation de $\tilde{Y}_k$ dans la méthode STATIS.

On peut aussi voir qu'il existe une certaine dualité entre les méthodes procustéennes et les modèles INDSCAL et IDIOSCAL:

- Pour ces derniers, on suppose que les individus occupent une position donnée, et que chaque juge évalue les distances entre individus

en fonction d'une métrique qui lui est propre .

- Pour le modèle Procuste,la métrique M est commune à toutes les études,mais la position des individus dans l'espace $\mathbb{R}^p$ est différente d'une étude à l'autre .

*

* *

CONCLUSION GENERALE :

Nous avons présenté plusieurs techniques d'analyse conjointe de plusieurs matrices de données,en montrant les liens existant entre elles . A part le modèle INDSCAL qui ne semble applicable que si les données ont une structure très particulière,ces méthodes s'appliquent à une très grande variété de situations,décrites au chapitre II .

Cependant,lorsqu'on dispose de données chronologiques(données obtenues à des instants différents),le temps n'intervient ici que comme un outil d'interprétation des résultats:on parle d'évolution des variables ou des individus au cours du temps;on ne tient pas comptedans l'analyse de la chronologie des observations.

On peut cependant citer le travail de M. TENENHAUS et B.PRIEURET(31), ainsi que la thèse de J.M. BOUROCHE(3),où le temps est utilisé de façon plus explicite.Mais cette approche est assez limitée:on suppose que le temps n'intervient que sur l'évolution des centres de gravité des tableaux.Il serait intéressant de chercher comment mieux exploiter des séries chronologiques multidimensionnelles par des méthodes généralisant les techniques classiques d'Analyse des Données.

D'autre part,nous n'avons pas envisagé l'étude de la situation 5 (mêmes variables,mêmes individus) du point de vue de l'analyse d'un cube de données,dont les trois arêtes joueraient des rôles similaires, par une généralisation de dimension 3 du schéma de dualité;cette approche,appelée "Analyse Triadique" par P.A. JAFFRENOU dans sa thèse (15), a été aussi abordée par L.R.TUCKER (30) .

*

* *

BIBLIOGRAPHIE

(1) M.C. BERNARD, F.J. DIAZ-LLANOS, Y. ESCOUFIER (1979) - "La méthode Statis : Une application à l'évolution des campagnes languedociennes". C.R.I.G., avenue d'Occitanie - 34075 Montpellier Cedex - Rapport Technique n° 7905.

(2) P. BOURGEOIS (1980) - "Recherche du déplacement minimisant la distance entre deux configurations de points indicés par un même ensemble fini. Méthodes et applications en reconnaissance des formes et en analyse des données cubiques". Thèse - Université Pierre et Marie Curie - Paris VI.

(3) J.M. BOUROCHE (1975) - "Analyse des données ternaires : La double analyse en composantes principales". Thèse - Université Pierre et Marie Curie - Paris VI.

(4) F. CAILLIEZ, J.P. PAGES (1976) - "Introduction à l'analyse des données". SMASH - 9, rue Duban, 75016 Paris.

(5) J.D. CARROLL (1968) - "A generalization of canonical analysis to three or more sets of variables". Prc. 76th Convention, American Psychology Association, p. 227-228.

(6) J.D. CARROLL, J.J. CHANG (1970) - "Analysis of individual differences in Multidimensional Scaling via an N-way generalization of "Eckart-Young" decomposition". Psychometrika, Vol. 35 n° 3, p. 283-320.

(7) P. CAZES , S. BONNEFOUS, A. BAUMERDER, J.P. PAGES (1976) - "Description cohérente des variables qualitatives prises globalement et de leurs modalités". SAD 2/1976, p. 48-62.

(8) P. CAZES, A. BAUMERDER, S. BONNEFOUS, J.P. PAGES (1977) - "Codage et analyse des tableaux logiques. Introduction à la pratique des variables qualitatives". Cahiers du BURO n° 27 - Université de Paris VI.

(9) B. CHARLES (1981) - "Approximation d'opérateurs au sens de la trace". Séminaire d'analyse des données - C.R.I.G. Montpellier.

(10) J. DE LEEUW, S. PRUZANSKY (1978) - "A new computational method to fit the weighted euclidean distance model". Psychometrika, Vol. 43 n° 4, p. 479- 490.

(11) Y. ESCOUFIER, J.P. PAGES, P. CAZES (1976) - "Opérateurs et analyse des tableaux à plus de deux dimensions". Cahiers du BURO, n° 25, p. 61-89.

(12) Y. ESCOUFIER (1979) - "Cours d'analyse des données".
C.R.I.G., av. d'Occitanie 34075 Montpellier Cedex.

(13) T. FOUCART (1979) - "Structure des tableaux de probabilités - Description et prévision".
Thèse - Université des Sciences et Techniques du Languedoc Montpellier II.

(14) J.C. GOWER (1975) - "Generalized Procrustes Analysis".
Psychometrika, Vol. 40 n° 1, p.33-51.

(15) P.A. JAFFRENOU (1978) - "Sur l'analyse des familles finies de variables vectorielles".
Thèse - Université Claude Bernard - Lyon I.

(16) J.R. KETTENRING (1971) - "Canonical Analysis of several sets of variables"
Biometrika, 58, 3, p. 433-451.

(17) B. KORTH, L.R. TUCKER (1976) - "Procrustes matching by congruence coefficients".
Psychometrika, 41, n° 4, p. 531-536.

(18) H. L'HERMIER DES PLANTES (1976) - "Structuration des tableaux à trois indices de la statistique".
Thèse - Université des Sciences et Techniques du Languedoc - Montpellier II.

(19) H. L'HERMIER DES PLANTES, Y. ESCOUFIER (1978) - "A propos de la comparaison graphique des matrices de variances".
Biom. Journal, Vol. 20 n° 5, p. 477-483.

(20) H. L'HERMIER DES PLANTES, B. THIEBAUT (1977) - "Etude de la pluviosité au moyen de la méthode Statis".

(21) J.C. LINGOES, P.H. SCHÖNEMANN (1974) - "Alternative measures of fit for the Schönemann - Carroll matrix fitting algorithm".
Psychometrika, Vol. 39 n° 4.

(22) J.P. MAILLES (1978) - "Analyse factorielle des tableaux de dissimilarités".
Thèse - Université Pierre et Marie Curie - Paris VI.

(23) M.C. PLACE (1980) - "Contribution algorithmique à la mise en oeuvre de la méthode Statis".
Thèse - Université des Sciences et Techniques du Languedoc - Montpellier II.

(24) C.R. RAO (1965) - "The use and interpretation of principal component analysis in applied research".
Sankhya A., 26, p. 329-358.

(25) P.H. SCHÖNEMANN (1972) - "An algebraic solution for a class of subjective metrics models".
Psychometrika, Vol. 37 n° 4, p. 441-451.

(26) P.H. SCHÖNEMANN, R.M. CARROLL (1970) - "Fitting one matrix to another under choice of a central dilatation and a rigid motion".
Psychometrika, Vol. 35 n° 2, p. 245-256.

(27) R. SIBSON (1978) - "Studies in the robustness of multidimensional scaling : Procrustes statistics".
J.R. Statis. Soc. B. Vol. 40 n° 2.

(28) J.M.F. TEN BERGE (1977) - "Orthogonal Procrustes rotation for two or more matrices".
Psychometrika, Vol. 42 n° 2.

(29) J.M.F. TEN BERGE, K. NEVELS (1977) -"A general solution to Mosier's oblique procrustes problem".
Psychometrika, Vol. 42 n° 4.

(30) L.R. TUCKER (1972) - "Relations beetween multidimensional scaling and three-mode factor analysis".
Psychometrika, Vol. 37 n° 1, p. 3-27.

(31) M. TENENHAUS, B. PRIEURET (1971) -"Analyse des séries chronologiques multidimensionnelles" .
Université d'Ottawa .

*

* *

Dernière page d'une thèse

VU

Grenoble, le 27/5/1987

Le Président de la thèse

Vu, et permis d'imprimer,

Grenoble, le 2.6.87

Le Président de l'Université Scientifique et Médicale

Le Président,

J.J. PAYAN

UNIVERSITÉ SCIENTIFIQUE ET MÉDICALE DE GRENOBLE