

# CAHIERS DU BURO

A. CROQUETTE

A. KOBILINSKY

J. P. PAGES

Y. SCHEKTMAN

**Optimalité en analyse factorielle Vers de nouveaux critères et de nouvelles stratégies. Deuxième partie : Quelques propriétés extrémales des directions, composantes et sous-espaces principaux, introduction aux analyses factorielles sous contraintes**

*Cahiers du Bureau universitaire de recherche opérationnelle.*  
*Série Recherche*, tome 46-47 (1986), p. 3-35

[http://www.numdam.org/item?id=BURO\\_1986\\_\\_46-47\\_\\_3\\_0](http://www.numdam.org/item?id=BURO_1986__46-47__3_0)

© Institut Henri Poincaré — Institut de statistique de l'université de Paris, 1986, tous droits réservés.

L'accès aux archives de la revue « Cahiers du Bureau universitaire de recherche opérationnelle. Série Recherche » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## OPTIMALITE EN ANALYSE FACTORIELLE

Vers de nouveaux critères et de nouvelles stratégies

A. CROQUETTE<sup>1</sup>, A. KOBILINSKY<sup>2</sup>, J.P. PAGES<sup>3</sup>, Y. SCHEKTMAN<sup>1</sup>

### DEUXIÈME PARTIE :

QUELQUES PROPRIETES EXTREMALES DES DIRECTIONS,  
COMPOSANTES ET SOUS-ESPACES PRINCIPAUX,  
INTRODUCTION AUX ANALYSES FACTORIELLES SOUS CONTRAINTES.

<sup>1</sup> *Laboratoire Modèles et logiciels d'analyse de données.  
Univ. Paul Sabatier. Toulouse - GRECO-CNRS 59.*

<sup>2</sup> *INRA-CNRA, Laboratoire de Biométrie.  
Route de St Cyr 78000 Versailles.*

<sup>3</sup> *Laboratoire de Statistique et d'Etudes Economiques et Sociales.  
Centre d'Etudes Nucléaires de Fontenay-aux-Roses.*



SOMMAIRE DE LA DEUXIEME PARTIE

	Pages
1. CRITERES BASES SUR LES VALEURS PROPRES	7
2. CRITERES BASES SUR DES QUOTIENTS DE RAYLEIGH	12
2.1 Présentation du problème	12
2.2 Le domaine des inerties associées à des directions orthogonales	14
2.3 Les fonctions adaptées aux critères	18
2.4 Typologie des propriétés extrémales	22
2.5 Quotients de Rayleigh et dualité	24
2.6 Remarque sur le domaine du spectre des opérateurs AVA'M	24
3. OPTIMALITE ET ANALYSE FACTORIELLE SOUS CONTRAINTES	26
3.1 Analyses orthogonales et obliques	27
3.2 Quelques analyses utiles - Optimalité sous contraintes	30
BIBLIOGRAPHIE	34

Soit  $E$  un espace vectoriel de dimension finie et  $V$  une forme quadratique définie-positive sur le dual  $E^*$  de  $E$  ; à cette forme quadratique on peut associer les métriques euclidiennes  $V$  et  $V^{-1}$  sur  $E^*$  et  $E$  respectivement. On sait que l'étude des propriétés d'optimalité des analyses factorielles revient à l'exploitation des décompositions optimales des formes quadratiques telles que  $V$  .

Dans la première partie de cet article [A. Croquette, A. Kobilinsky, J.P. Pagès, Y. Schektman - 1984], la recherche des ajustements par en dessous de  $V$  a reposé uniquement sur les inégalités usuelles entre formes quadratiques.

L'exploitation du préordre  $O_1$  (il correspond dans  $E$  à l'inclusion des ellipsoïdes de concentration) n'a fait intervenir qu'une seule métrique sur  $E$  : la métrique  $V^{-1}$  . La classe des formes quadratiques sous dominantes de rang  $q$  au sens de  $O_1$  coïncide avec la classe de  $AVA'$  où l'opérateur  $A$  parcourt l'ensemble des projecteurs  $V^{-1}$ -symétriques.

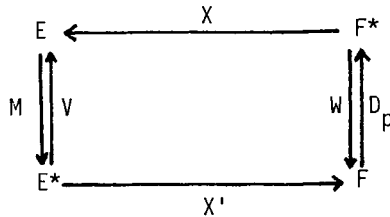
En introduisant avec le  $M$ -préordre  $O_2$  (comparaison, dans la métrique  $M$  , des longueurs des axes des ellipsoïdes) une deuxième métrique sur  $E$  , on a pu faire un choix parmi ces formes quadratiques. Le projecteur  $V^{-1}$ -symétrique  $A$  , qui fournit l'ajustement sous dominant  $AVA'$  de rang  $q$  , maximum pour le  $M$ -préordre  $O_2$  , est aussi  $M$ -symétrique ; son image n'est autre que le sous-espace principal  $E_q$  de dimension  $q$  .

La deuxième partie de cet article est consacrée aux critères d'optimalité qui permettent de retrouver les axes, composantes et sous-espaces principaux et à l'application de ces critères aux analyses factorielles sous contraintes.

Dans un premier chapitre on rappelle brièvement quels sont les critères numériques caractérisant  $E_q$  qui découlent immédiatement des résultats présentés dans la première partie. Voulant échapper aux fonctions des valeurs propres, on développe ensuite dans un deuxième chapitre une autre famille de critères faisant intervenir la notion générale de quotient de Rayleigh. Ayant montré le lien entre les deux approches, on est alors à même d'aborder dans un troisième chapitre d'autres stratégies

dans la recherche d'axes ou de sous-espaces principaux.

Comme dans la première partie, au tableau de données centré  $X$  et aux deux métriques euclidiennes  $M$  et  $D_p$  (métrique des poids) qui permettent de mesurer les proximités entre ses  $n$  colonnes et ses  $p$  lignes respectivement est associé le triplet  $(X, M, D_p)$  et le schéma de dualité :



avec :  $V = X D_p X'$  ;  $W = X' M X$  .

### 1. CRITERES BASES SUR LES VALEURS PROPRES

Considérons la décomposition :

$$V = V_1 + V_2 .$$

Jusqu'à présent les décompositions optimales envisagées conduisaient à des formes  $V_1$  et  $V_2$  de type  $AVA'$  , où  $A$  était un projecteur soit  $V^{-1}$ -symétrique (préordre  $O_1$ ) soit  $M$  et  $V^{-1}$ -symétrique ( $M$ -préordre  $O_2$ ). Ces décompositions avaient des interprétations géométriques simples.

Dès que l'on introduit des contraintes les décompositions que l'on obtient ne possèdent plus ces "bonnes propriétés" ; des contradictions apparaissent et il devient difficile d'avoir recours à l'intuition pour trouver les ajustements optimaux.

Considérons, par exemple, la décomposition de la forme quadratique  $V$  obtenue à partir d'un ajustement par en-dessous de rang  $q$  d'une forme  $W \ll V$  maximal au sens du même préordre  $O_2$  . Dans cette décomposition :  $V = V_1 + V_2$  , la forme quadratique  $V_1$  s'écrit :  $V_1 = AWA'$  , où  $A$  est un projecteur  $M$  et  $W^{-1}$ -symétrique. Manifestement  $V_1$  est aussi un ajustement

tement par en-dessous de  $V$ , vérifiant la contrainte :  $V_1 \prec W$ , maximal au sens du M-préordre  $O_2$ . Cet ajustement conduit au résidu :

$$V_2 = V - V_1 = V - W + W - V_1 .$$

Ce résidu ne vérifie pas les bonnes propriétés : il n'est pas du type  $AVA'$ , ni du type  $AWA'$  ; il est en général de rang supérieur à  $p-q$ . Allons plus loin. Soit  $V'_1$  un autre ajustement par en-dessous de  $V$ , de rang  $q$  et vérifiant la contrainte :  $V'_1 \prec W$ . La forme  $V'_1$ , étant issue d'une décomposition de  $W$ , est inférieure à  $V_1$  au sens de  $O_2$ . Si la comparaison entre les différents ajustements sous contrainte de rang  $q$  de  $V$  est donc possible, elle ne peut s'effectuer au niveau des résidus. En effet, si  $V'_1$  est bien inférieur à  $V_1$  au sens de  $O_2$ , le résidu  $V'_2 = V - V'_1$  peut ne pas être comparable à  $V_2 = V - V_1$  au sens de ce préordre. C'est parce que l'ordre sur les spectres n'est pas toujours conservé quand on rajoute une même forme quadratique (ici :  $V - W$ ), que la contradiction précédente apparaît entre analyse des ajustements et analyse des résidus.

Si le recours aux M-préordres  $O_2$  et  $O'_2$  ( $O'_2$  généralise  $O_2$  dans le cas de formes quadratiques quelconques) pose problème dès que l'on opère sous contrainte, il n'en est plus de même quand on utilise des critères qui font appel à des fonctions du spectre des opérateurs considérés ; ces fonctions induisent des préordres totaux sur des formes quadratiques plus grossiers que les préordres  $O_1$  et  $O_2$ . C'est à l'étude de ces fonctions et des critères qui leurs sont associés que ce chapitre est consacré.

### Critères sur les variables

Notons  $\lambda_1(A), \lambda_2(A), \dots, \lambda_p(A)$  les valeurs propres, rangées par valeurs décroissantes, d'un opérateur diagonalisable  $A$  défini sur  $E = \mathbb{R}^p$ . Soit  $f(x_1, \dots, x_p)$  une fonction numérique sur  $E$  croissante en chaque argument.

Définition 1 : (f,M)-préordre  $0_3$

$V_1$  est dite (f,M)-inférieur à  $V_2$  si on a l'inégalité :

$$f[\lambda_1(V_1M), \dots, \lambda_p(V_1M)] < f[\lambda_1(V_2M), \dots, \lambda_p(V_2M)]$$

on écrit alors :  $V_1 \underset{(f,M)}{<} V_2$  .

Il est clair que le (f,M)-préordre  $0_3$  est plus grossier que le M-préordre  $0_2$  :

$$\text{si } V_1 \underset{M}{<} V_2 : V_1 \underset{(f,M)}{<} V_2$$

Par suite, un élément maximal (respectivement minimal) au sens de  $0_2$  l'est aussi au sens de  $0_3$  . En outre, si f est strictement croissante en chaque argument, l'unicité pour  $0_2$  de l'élément maximal garantit son unicité pour  $0_3$  . Cette dernière condition peut dans certains cas être affaiblie en croissance stricte vis à vis des q premiers arguments, sur le domaine où ceux-ci sont tous strictement positifs.

Sachant que l'on a noté  $(X, M, D_p)$  le triplet associé au tableau de données centré X à p lignes (caractères  $x^j$ ) et n colonnes (individus  $x_i$ ) et aux deux métriques euclidiennes M et  $D_p$  définies respectivement sur  $E = R^p$  et  $F = R^n$  , rappelons les deux exemples les plus classiques de fonction f .

Exemple 1 : le critère de la trace

$$\text{si } : f(\lambda_1, \dots, \lambda_p) = \lambda_1 + \dots + \lambda_p \quad ; \quad \text{on a } :$$

$$f[\lambda_1(V_1M), \dots, \lambda_p(V_1M)] = \text{tr}(V_1M) \quad .$$

La matrice  $V_1$ , étant inférieure à V au sens de  $0_1$  , peut être considérée comme la matrice de variance expliquée par un sous-espace  $F_1$  de F de dimension q ; on peut donc écrire :

$$V_1 = X A' D_p A X'$$



où  $A$  est l'opérateur de  $D_p$ -projection sur  $F_1$ .

Si les variables  $x^j$  sont réduites et si  $M = I_p$  (métrique identité), le critère pour le sous-espace  $F_1$  s'écrit :

$$\text{tr}(V_1 M) = \text{tr}(X A' D_p A X') = \sum_{j=1}^p D_p(x^j, A x^j) = \sum_{j=1}^p \cos^2 \theta_j$$

où  $\cos \theta_j$  désigne le coefficient de corrélation multiple entre la variable  $x^j$  et le sous-espace  $F_1$ .

On retrouve donc ici, dans le cas où  $q=1$ , le critère introduit par H. Hotelling en 1933. Si les variables  $c^1, \dots, c^q$  forment une base orthogonale de  $F_1$ , le critère s'écrit encore :

$$\text{tr}(V_1 M) = \sum_{j=1}^q \frac{D_p(W D_p c^j, c^j)}{D_p(c^j, c^j)} ;$$

il apparaît ici comme une somme de "quotients de Rayleigh".

Exemple 2 : le critère du déterminant

$$\text{si } f(\lambda_1, \dots, \lambda_p) = \lambda_1 \dots \lambda_q ,$$

$f(\lambda_1(V_1 M), \dots, \lambda_p(V_1 M))$  est le déterminant de l'opérateur  $V_1 M$  restreint à l'image de dimension  $q$  de  $V_1$ .

Ce déterminant est maximum si et seulement si  $V_1 = A V A'$  où  $A$  est le projecteur  $V^{-1}$ -symétrique sur l'un des sous-espaces principaux de dimension  $q$ .

Si on utilise le même critère pour rechercher le minimum de  $V-W$ , où  $W$  varie parmi les formes quadratiques de rang  $p-q$  inférieures à  $V$ , on ne retombe pas sur le même résultat. En effet, considérons par exemple le cas où :

$$p=2, \quad q=1, \quad V = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad M = I .$$

La matrice  $W = \begin{pmatrix} 1,5 & 0 \\ 0 & 0 \end{pmatrix}$  est aussi proche de  $V$  au sens du critère du déterminant que la matrice  $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$  qui est l'ajustement optimal au sens du  $M$ -préordre  $0_2$ .

Au préordre  $0_2$ , on a associé des fonctions du spectre croissantes en chaque argument. De même au préordre  $0_2'$ , qui a permis de trouver la forme  $W$  quelconque la plus "proche" de  $V$ , on peut associer toute fonction du spectre dont la variation en fonction d'un argument est croissante si cet argument est positif et décroissante dans le cas contraire. Un cas particulier bien connu de telle fonction est la somme des carrés des valeurs propres de  $(V-W)M$ , égale à la trace du carré de cet opérateur. Elle définit une distance entre  $V$  et  $W$  qui est donc minimum lorsque  $W = AVA'$ , où  $A$  est le projecteur  $V^{-1}$ -symétrique sur le sous-espace propre associé aux  $p-q$  dernières valeurs propres de  $VM$ .

### Critère sur les individus

L'approche factorielle cherche à reconstituer la matrice de variance  $V$  au moyen d'un nombre réduit  $q$  de facteurs. Une approche symétrique peut être faite sur les individus. Le problème est alors de reconstituer au mieux les distances entre individus, ou de façon équivalente les produits scalaires donné par la matrice  $W = X'MX$ , à partir des projections  $M$ -orthogonales sur un sous-espace  $E_1$  de dimension  $q$ .

Le critère de la trace, comme précédemment, peut être retenu ; si les variables  $u_i$  forment une base orthogonale du sous-espace, de dimension  $q$ ,  $E_1$  de  $E$ , ce critère s'écrit comme une somme de quotients de Rayleigh :

$$\text{tr}(W_1 D_p) = \sum_{i=1}^q \frac{N(VMu_i, u_i)}{N(u_i, u_i)} .$$

## 2. CRITERES BASES SUR DES QUOTIENTS DE RAYLEIGH

Les critères basés sur les valeurs propres du chapitre précédent mettent en évidence une propriété d'optimalité des directions principales dans la classe des directions propres d'une famille d'opérateurs (les ajustements sous dominants  $AVA'$  de  $V$ ). Cette classe est trop contraignante dans le contexte des Analyses en Composantes Principales sous Contraintes éventuellement non linéaires ; d'autres propriétés extrémales des directions principales, dans la classe des directions orthogonales pour une métrique fixée a priori, ont été développées [Schektman 1978]. Ce chapitre étudie de façon complète une famille de critères relevant de cette démarche : la famille de critères basés sur la notion générale de quotient de Rayleigh. Les premiers résultats sont établis dans le cas particulier où ces quotients s'identifient à des inerties ; cette approche permet d'éclairer les liens entre ces critères et ceux basés sur les valeurs propres.

### 2.1 Présentation du problème

Les directions principales  $\Delta u_i$ , qu'usuellement on obtient en diagonalisant l'opérateur  $VM$ , peuvent être présentées comme les solutions d'un problème d'optimisation d'une fonction numérique de directions orthogonales au sens de la métrique  $M$ . Illustrons cette démarche en analysant les problèmes  $R_q$ ,  $S_q$  et  $S'_q$  introduits respectivement par [RAO (1964)] et [SCHEKTMAN (1978)] pour lesquels on note :

$$I_{\Delta_u^\perp} = \frac{M(VMu, u)}{M(u, u)}$$

le moment d'inertie du nuage des points individus dans leur espace  $(E, M)$  par rapport à l'hyperplan  $\Delta_u^\perp$ , orthogonal à l'axe  $\Delta_u$  engendré par le vecteur  $u$ . Cette inertie mesure la dispersion de la  $M$ -projection sur l'axe  $\Delta_u$  du nuage considéré ; elle est exprimée ci-dessus sous la forme d'un quotient de Rayleigh.

$$\begin{array}{l}
 R_q : \left\{ \begin{array}{l} \text{MAXIMISER} \quad \sum_{i=1}^q I_{\Delta_{u_i}^\perp} \\ \text{sous les contraintes : } N(u_i, u_{i'}) = 0 \quad \text{si } i \neq i' \end{array} \right. \\
 \\
 S_q : \left\{ \begin{array}{l} \text{MAXIMISER} \quad \sum_{i=1}^q [ I_{\Delta_{u_i}^\perp} ]^2 \\ \text{sous les contraintes : } M(u_i, u_{i'}) = 0 \quad \text{si } i \neq i' \end{array} \right. \\
 \\
 S_q' : \left\{ \begin{array}{l} \text{MINIMISER} \quad \sum_{i=1}^q I_{\Delta_{u_i}^\perp} \\ \text{sous les contraintes : } M(u_i, u_{i'}) = 0 \quad \text{si } i \neq i' \end{array} \right.
 \end{array}$$

Dans le problème  $R_q$ , la somme des inerties (fonction objectif) est maximum pour, et seulement pour, tout système de directions orthogonales  $\Delta_{u_i}$  engendrant le sous-espace principal de dimension  $q$ ; la détermination des  $q$  premiers axes principaux nécessite la résolution pas à pas des problèmes  $R_1, R_2, \dots, R_q$ . Le problème  $S_q$  (resp.  $S_q'$ ) fournit directement à l'optimum les  $q$  premières (resp. dernières) directions principales. Ces remarques nous conduisent à poser la définition suivante :

Définition 2 : Critère sous-espace - Critère directionnel

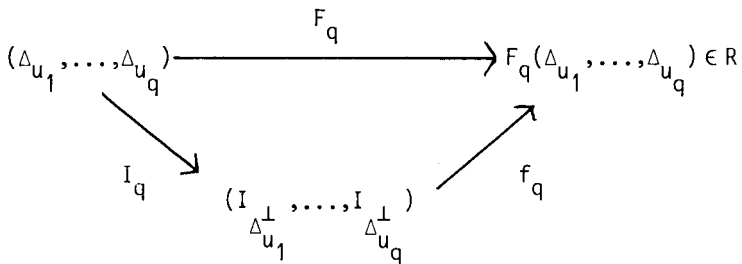
En analyse en Composantes Principales, on appelle :

- critère directionnel, tout critère permettant de retrouver les directions principales.
- critère sous-espace, tout critère qui, de façon plus faible, ne permet que de retrouver un sous-espace de  $E = \mathbb{R}^P$ , engendré par des directions principales.

L'objet de cette partie est de faire la synthèse des problèmes du type  $R_q$ ,  $S_q$  ou  $S'_q$  et de mettre en évidence des propriétés simples permettant de construire ou différencier les critères sous-espace et directionnels.

Remarquons d'abord que les fonctions objectifs  $F_q$  des problèmes d'optimisation considérés sont des fonctions d'axes  $\Delta_u$  orthogonaux par l'intermédiaire des inerties  $I_{\Delta_u^\perp}$  associées ; on a le schéma de factorisation

$F_q = f_q \circ I_q$  suivant :



$I_q$  associe à  $q$  directions orthogonales  $\Delta_{u_i}$  les moments d'inertie  $I_{\Delta_{u_i}^\perp}$  (quotients de Rayleigh) ; pour  $R_q$ ,  $f_q$  correspond à la somme des arguments, tandis que pour  $S_q$  et  $S'_q$ , elle correspond respectivement à la somme de leurs carrés et à leur produit. Notons également que les fonctions  $F_q$  ou  $f_q$  des problèmes d'optimisation précédents sont symétriques (invariantes par permutation de leurs arguments). Avant de proposer des familles de fonctions symétriques  $f_q$ , donnant naissance soit à des critères directionnels soit à des critères sous-espace, précisons le domaine des systèmes de quotients de Rayleigh, image de l'application  $I_q$ , sur lequel opèrent ces fonctions.

## 2.2 Le domaine des inerties associées à des directions orthogonales

Des conditions simples, vérifiées par un système d'inerties  $(I_{\Delta_{u_1}^\perp}, I_{\Delta_{u_2}^\perp}, \dots, I_{\Delta_{u_q}^\perp})$  associées à  $q$  directions  $\Delta_{u_i}$  orthogonales, sont

suggérées par le critère de RAO ; on a,  $\lambda_1 > \lambda_2 > \dots > \lambda_p$  désignant les

moments principaux de l'Analyse en Composantes Principales du triplet  $(X, M, D_p)$  :

$$(1) \quad \sum_{i=1}^q I_{\Delta_{U_i}^\perp} < \sum_{i=1}^q \lambda_i \quad \text{pour } q = 1, 2, \dots, p$$

Notons qu'on a aussi :

$$(1) \quad \sum_{i=1}^q \lambda_{p-i+1} < \sum_{i=1}^q I_{\Delta_{U_i}^\perp} \quad \text{pour } q = 1, 2, \dots, p$$

et que toutes ces sommes sont égales pour  $q=p$  à l'inertie trace (VM) du nuage des individus de  $E$ . La proposition 1 montre que l'ensemble des  $q$  inerties (système des quotients de Rayleigh), image de l'application  $I_q$ , est caractérisé par un système d'inégalités de ce type.

PROPOSITION 1 : CARACTERISATION ANALYTIQUE DU DOMAINE DES INERTIES

L'ensemble des  $q$ -uplets d'inerties du nuage des points individus de  $E = \mathbb{R}^p$  associées à  $q$  directions orthogonales est l'ensemble convexe  $C_q(\lambda_1, \lambda_2, \dots, \lambda_p)$  des points  $(x_1, x_2, \dots, x_q)$  de  $\mathbb{R}^q$  tels que :

$$(2) \quad \sum_{i=1}^{|J|} \lambda_{p-i+1} < \sum_{i \in J} x_i < \sum_{i=1}^{|J|} \lambda_i$$

pour toute partie  $J$  de  $\{1, 2, \dots, q\}$  de cardinal  $|J|$  inférieur ou égal à  $q$ .

La démonstration de la proposition 1 repose sur le lemme suivant pour lequel les  $x_i$  et les  $\lambda_i$  ne font pas a priori référence à des inerties et à des valeurs propres.

LEMME 1 : Pour tout point  $(x_1, x_2, \dots, x_{q-1}, x_q)$  de  $C_q(\lambda_1, \lambda_2, \dots, \lambda_p)$ ,  $(x_1, x_2, \dots, x_{q-1})$  appartient à l'ensemble  $C_{q-1}(\lambda'_1, \lambda'_2, \dots, \lambda'_{p-1})$  où

$$\lambda'_i = \begin{cases} \lambda_i & \text{pour } i < \ell \\ \lambda_\ell + \lambda_{\ell+1} - x_q & \text{pour } i = \ell \\ \lambda_{i+1} & \text{pour } i > \ell \end{cases}$$

$\ell$  étant tel que  $\lambda_{\ell+1} < x_q < \lambda_\ell$  .

Démonstration :

Pour  $J = \{q\}$ , (2) s'écrit  $\lambda_p < x_q < \lambda_1$  ; il existe donc  $\ell$  tel que  $\lambda_{\ell+1} < x_q < \lambda_\ell$  ; d'autre part, pour toute partie  $J$  ne contenant pas  $q$  telle que  $|J| < q-1$ , les inégalités (2) appliquées à  $J$  et  $J \cup \{q\}$  conduisent à :

$$\text{MAX} \left[ \sum_{i=1}^{|J|} \lambda_{p-i+1} ; \sum_{i=1}^{|J|+1} \lambda_{p-i+1} - x_q \right] < \sum_{i \in J} x_i < \text{MIN} \left[ \sum_{i=1}^{|J|} \lambda_i ; \sum_{i=1}^{|J|+1} \lambda_i - x_q \right]$$

Elles s'écrivent en fonction des nombres  $\lambda'_i$  :

$$\sum_{i=1}^{|J|} \lambda_{(p-1)-i+1} < \sum_{i \in J} x_i < \sum_{i=1}^{|J|} \lambda'_i$$

et traduisent l'appartenance de  $(x_1, x_2, \dots, x_{q-1})$  à  $C_{q-1}(\lambda'_1, \lambda'_2, \dots, \lambda'_{p-1})$ .

Démonstration de la proposition 1 :

D'après (1), le système des inerties  $I_{\Delta u_i}$  vérifie (2) ; ceci s'exprime encore ainsi : l'image de  $I_q$  est dans  $C_q(\lambda_1, \lambda_2, \dots, \lambda_p)$ . Montrons que, réciproquement,  $C_q$  contient l'image de  $I_q$  .

Soit  $(x_1, x_2, \dots, x_q)$  un point de  $C_q$  ; avec les notations du lemme 1, définissons le vecteur  $u_q$  en fonction des vecteurs axiaux principaux  $u_\ell$  et  $u_{\ell+1}$  par :

$$u_q = \begin{cases} \sqrt{x_q - \lambda_{\ell+1}} u_\ell + \sqrt{\lambda_\ell - x_q} u_{\ell+1} & \text{si } \lambda_\ell \neq \lambda_{\ell+1} \\ u_\ell & \text{si } \lambda_\ell = \lambda_{\ell+1} \end{cases}$$

$$\text{On a : } x_q = I_{\Delta_{u_q}^\perp} \left( = \frac{M(VMu_q, u_q)}{M(u_q, u_q)} \right)$$

et d'après le lemme 1 :  $(x_1, x_2, \dots, x_{q-1}) \in C_{q-1}(\lambda'_1, \lambda'_2, \dots, \lambda'_{p-1})$ .

Remarquons que, si B est le projecteur M-symétrique de E sur l'hyperplan  $\Delta_{u_q}^\perp$  orthogonal à  $\Delta_{u_q}$ , les nombres  $\lambda'_1 > \lambda'_2 > \dots > \lambda'_{p-1}$ ,

définis dans le lemme 1, sont les valeurs propres de la restriction de BVM à l'image de B. Le raisonnement précédent, effectué alors en remplaçant dans un premier temps VM par cette restriction, conduit à un deuxième axe  $\Delta_{u_{q-1}}^\perp$  tel que  $x_{q-1} = I_{\Delta_{u_{q-1}}^\perp}$ . En recommençant ce type

d'opération autant de fois que nécessaire, on aboutit aux axes  $\Delta_{u_q}, \Delta_{u_{q-1}}, \dots, \Delta_{u_1}$ , orthogonaux par construction ; les coordonnées

$x_i$  sont les moments d'inertie  $I_{\Delta_{u_i}^\perp}$  et  $(x_1, x_2, \dots, x_q)$  est un point de

l'image de  $I_q$ .

Indiquons quelques propriétés des domaines d'inertie  $C_q$  qui permettent de mieux les appréhender. On peut montrer [Croquette - 1980] que le polyèdre convexe  $C_q$  s'identifie à la projection cartésienne de  $C_p$  sur  $\mathbb{R}^q$ . De plus le domaine  $C_q$  respecte certaines symétries : tout point obtenu à partir d'un point de  $C_q$  en permutant ses coordonnées appartient à  $C_q$ . Enfin  $C_q$  est l'enveloppe convexe symétrique des points (extrémaux) obtenus en supprimant dans le spectre ordonné  $(\lambda_1, \lambda_2, \dots, \lambda_p)$  de VM,  $p-q$  valeurs propres successives. La figure 1 illustre dans des cas simples ces domaines : les points  $\Lambda_{j..q}$  et  $\bar{\Lambda}$  désignent respectivement  $(\lambda_j, \dots, \lambda_q)$  et  $(\bar{\lambda}, \bar{\lambda}, \dots, \bar{\lambda})$  où  $\bar{\lambda}$  est la moyenne des  $p$  valeurs propres  $\lambda_j$ .



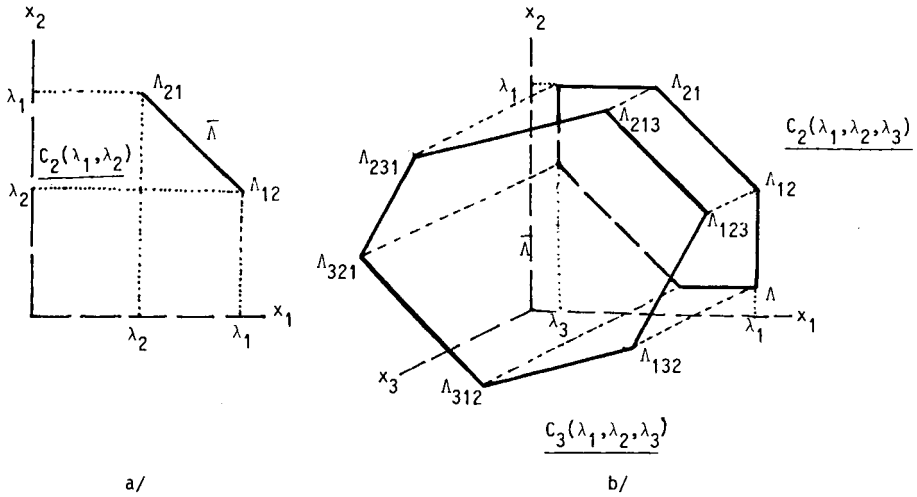


Figure 1

Domaine des systèmes d'inertie (quotients de Rayleigh) associées à des axes orthogonaux (a/ pour  $p = q = 2$  ; b/ pour  $p = 3, q = 1, 2, 3$ ).

### 2.3 Les fonctions adaptées aux critères

Pour rechercher les fonctions objectif  $F_q$  symétriques permettant de retrouver les  $q$  premiers axes principaux, utilisons la factorisation  $F_q = f_q \circ I_q$  du paragraphe 2.1 et la connaissance du domaine des systèmes d'inertie, image de l'application  $I_q$ . Le support géométrique de la figure 1a/ conduit, dans le cas  $p = q = 2$ , aux constatations suivantes :  $F_2$  est maximum pour les directions principales si et seulement si la fonction symétrique  $f_2$ , définie sur  $C_2(\lambda_1, \lambda_2)$ , est maximum en  $\Lambda_{12} = (\lambda_1, \lambda_2)$  ; ceci est réalisé en imposant à  $f_2$  de croître en s'éloignant de  $\bar{\Lambda}$ , c'est-à-dire "en s'éloignant orthogonalement de la droite d'équation  $x_1 = x_2$ ", l'unicité du maximum étant assurée grâce à une croissance stricte de  $f_2$ . Il s'avère que cette double propriété caractérise de façon géométrique et opérationnelle les fonctions

S-convexes (convexes au sens de SCHUR) introduites par Ostrowski en 1952 dans un contexte différent. La définition suivante permet de ramener l'étude de la S-convexité d'une fonction à l'étude de la croissance d'une fonction numérique d'une seule variable sur l'intervalle  $[0,1]$  .

DEFINITION 3 : S-convexité

Soit  $C$  un domaine convexe de  $\mathbb{R}^q$  symétrique par rapport aux hyperplans définis par les égalités :  $x_i = x_j$  ( $i < j$ ) .

La fonction  $f : C \rightarrow \mathbb{R}$  est dite (strictement) S-convexe dans  $C$  si et seulement si :

(i)  $f$  est symétrique, c'est-à-dire invariante par permutation de ses arguments.

(ii)  $f$  est (strictement) croissante en "s'éloignant orthogonalement de l'hyperplan d'équation  $x_1 = x_2$ ".

Si  $-f$  est S-convexe,  $f$  est dite S-concave.

Précisons (ii) : Soit  $x = (x_1, x_2, x_3, \dots, x_q) \in C$  et

$$x(t) = \left( \frac{x_1+x_2}{2} + t \frac{x_1-x_2}{2}, \frac{x_2+x_1}{2} + t \frac{x_2-x_1}{2}, x_3, \dots, x_q \right),$$
 la condition

(ii) peut s'expliciter sous la forme suivante : quel que soit  $x$  de  $C$  tel que  $x_1 \neq x_2$  , l'application  $t \rightarrow f[x(t)]$  est (strictement) croissante sur  $[0,1]$ .

Notons que, puisque  $f$  est symétrique, on peut remplacer dans cette définition le couple de coordonnées  $(x_1, x_2)$  par tout autre couple. Sur la figure 2, l'orientation des flèches indique le sens de croissance (stricte) d'une fonction (strictement) S-convexe définie sur le domaine des systèmes d'inerties  $C_3(\lambda_1, \lambda_2, \lambda_3)$  associées à 3 directions orthogonales.

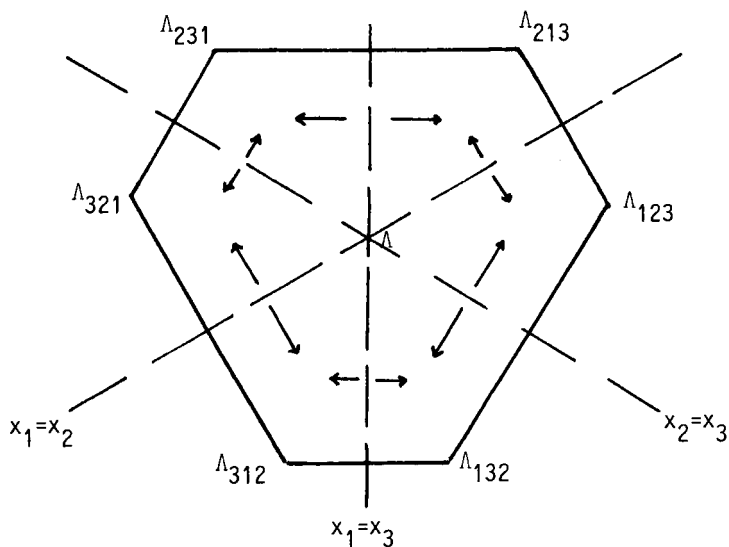


Figure 2

Symétries de  $C_3$  (pour  $p=3$ ) et directions de (stricte) croissance des fonctions (strictement)S-convexes sur  $C_3$  .

Nous allons montrer (théorème 1) qu'en circulant parallèlement aux arêtes du domaine  $C_q(\lambda_1, \lambda_2, \dots, \lambda_p)$  des systèmes d'inertie on atteint l'optimum ; l'ensemble des "fonctions adaptées" coïncidera alors avec l'ensemble  $\mathcal{F}_q(C_q)$  des fonctions de  $C_q$  dans  $R$

- strictement S-convexes si  $q=p$
- strictement S-convexes et strictement croissantes en chaque argument si  $q < p$  .

Le lemme 2 envisage le problème sur une arête de  $C_q$  .

LEMME 2 : Pour tout point  $(x_1, x_2, \dots, x_q)$  du domaine  $C_q(\lambda_1, \lambda_2, \dots, \lambda_p)$  des systèmes d'inertie et toute fonction  $f_q$  de  $\mathcal{F}_q(C_q)$  on a :

$$(3) \quad f_q(\lambda_1, \lambda_2, \dots, \lambda_{q-1}, \lambda_q) > f_q(\lambda_1', \lambda_2', \dots, \lambda_{q-1}', x_q)$$

où les nombres  $\lambda_j'$  sont définis comme dans le lemme 1 par

$$\lambda_i' = \begin{cases} \lambda_i & \text{pour } i < \ell \\ \lambda_\ell + \lambda_{\ell+1} - x_q & \text{pour } i = \ell \\ \lambda_{i+1} & \text{pour } i > \ell \end{cases}$$

$\ell$  étant tel que  $\lambda_{\ell+1} < x_q < \lambda_\ell$  .

De plus l'égalité a lieu si et seulement si à une permutation près les nombres  $\lambda_1', \lambda_2', \dots, \lambda_{q-1}', x_q$  coïncident avec les valeurs propres  $\lambda_1, \lambda_2, \dots, \lambda_q$  .

Montrons ce lemme pour  $q=p$  :

Si  $q=p$ , toutes les valeurs propres  $\lambda_i$  sont égales à un  $\lambda_j'$  sauf  $\lambda_\ell$  et  $\lambda_{\ell+1}$  qui vérifient  $\lambda_\ell + \lambda_{\ell+1} = \lambda_\ell' + x_p$  . Le point  $(\lambda_1', \lambda_2', \dots, \lambda_{p-1}', x_p)$  est donc situé sur l'une des arêtes de  $C_p$  . On atteint l'un des sommets de  $C_p$  , c'est-à-dire à une permutation près le point  $(\lambda_1, \lambda_2, \dots, \lambda_p)$  , "en s'éloignant orthogonalement de l'hyperplan  $x_\ell = x_{\ell+1}$ " , ce qui revient à rechercher l'optimum d'une fonction de  $\mathcal{F}_p(C_p)$ , ensemble des fonctions strictement S-convexes sur  $C_p$  .

THEOREME 1 :

Si  $f_q \in \mathcal{F}_q(C_q)$  alors la fonction objectif  $F_q = f_q \circ I_q$  est maximum pour et seulement pour le <sup>(\*)</sup> système des  $q$  premières directions propres (axes principaux)  $\Delta_{u_i}$  . Le maximum est égal à  $f_q(\lambda_1, \lambda_2, \dots, \lambda_q)$  .

---

(\*) "tout système orthogonal constitué de  $q$  premières.." si VM a des valeurs propres  $\lambda_i$  multiples.

Démonstration :

Soit  $(x_1, x_2, \dots, x_q) \in C_q$ , le système des inerties  $x_i = I_{\Delta v_i}^\perp$  associé aux directions  $(\Delta v_1, \Delta v_2, \dots, \Delta v_q)$  orthogonales de  $E$ .  
On a d'après le lemme 2 :

$$f_q(\lambda_1, \lambda_2, \dots, \lambda_q) \geq f_q(\lambda'_1, \lambda'_2, \dots, \lambda'_{q-1}, x_q)$$

et d'après le lemme 1 :

$$(x_1, x_2, \dots, x_{q-1}) \in C_{q-1}(\lambda'_1, \lambda'_2, \dots, \lambda'_{q-1}) .$$

Fixons  $x_q$  et posons  $f'_{q-1}(x_1, \dots, x_{q-1}) = f_q(x_1, \dots, x_{q-1}, x_q)$  ; en remplaçant  $(x_1, \dots, x_{q-1}, x_q)$  par  $(x_1, \dots, x_{q-1})$ ,  $f_q$  par  $f'_{q-1}$ ,  $C_q(\lambda_1, \dots, \lambda_{p-1}, \lambda_p)$  par  $C_{q-1}(\lambda'_1, \dots, \lambda'_{p-1})$ ,  $\mathcal{I}_q(C_q)$  par  $\mathcal{I}_{q-1}(C_{q-1})$  et en recommençant cette opération autant de fois que nécessaire, on aboutit à :

$$f_q(\lambda_1, \lambda_2, \dots, \lambda_q) \geq f_q(x_1, x_2, \dots, x_q)$$

$$\text{soit } f_q \circ I_q(u_1, u_2, \dots, u_q) \geq f_q \circ I_q(v_1, v_2, \dots, v_q) .$$

De plus l'égalité a lieu (lemme 2) si et seulement si à une permutation près  $(\lambda_1, \lambda_2, \dots, \lambda_q) = (x_1, x_2, \dots, x_q)$ ; soit si et seulement si :  
 $(\Delta v_1, \Delta v_2, \dots, \Delta v_q) = (\Delta u_1, \Delta u_2, \dots, \Delta u_q)$  .

En résumé, partant de l'optimum  $(\lambda_1, \lambda_2, \dots, \lambda_q)$ , on a montré que par un chemin en zig-zag dont les segments sont parallèles aux arêtes on peut aboutir en un point quelconque  $(x_1, x_2, \dots, x_q)$  de  $C_q$ . Les fonctions adaptées sont celles qui sont isotones sur ce chemin.

#### 2.4 Typologie des propriétés extrémales

Rappelons que les fonctions adaptées sont les fonctions  $f_q$  du domaine des systèmes d'inertie  $C_q$  dans  $R$  :

- strictement S-convexes si  $q = p$  .
- strictement S-convexes et strictement croissantes en chaque argument si  $q < p$  .

Si on remplace stricte S-convexité par S-convexité (large) le maximum de la fonction objectif  $F_q = f_q \circ I_q$  peut-être atteint pour des systèmes d'axes orthogonaux non principaux engendrant le sous-espace principal de dimension  $q$  . Dans ce contexte, qui est celui du critère de RAD  $R_q$  , seul le sous-espace principal de dimension  $q$  (et non les directions principales) est défini. On obtient ainsi un critère sous-espace au sens de la définition 2 .

Le tableau 1 résume les critères directionnels et sous-espace basés sur des quotients de Rayleigh. Dans ce tableau à double entrée les caractéristiques de chaque classe de critères sont précisées dans les cases correspondantes : on utilisera D (resp. E) pour indiquer que l'extremum (MIN ou MAX) est caractéristique des directions propres de B (resp. du sous-espace qu'elles engendrent), G si toutes les directions sont définies,  $P(m_1, m_2, \dots, m_q)$  si les seules directions définies sont celles dont les numéros sont dans la liste entre parenthèses.

Propriétés de $f$ sur $C_q$		Croissance stricte	Décroissance stricte	Autre
S-convexité	stricte	DP(1, ..., q) MAX	DP(p, ..., p-q-1) MAX	DG MAX
	large	EP(1, ..., q) MAX	EP(p, ..., p-q+1) MAX	EG MAX
S-concavité	stricte	DP(p, ..., p-q+1) MIN	DP(1, ..., q) MIN	DG MIN
	large	EP(p, ..., p-q+1) MIN	EP(1, ..., q) MIN	EG MIN

Tableau 1

Typologie des critères basés sur des quotients de Rayleigh associés à des directions orthogonales.

### 2.5 Quotients de Rayleigh et dualité

Les résultats précédents posent pour leur démonstration sur l'expression  $\frac{M(VMu, u)}{M(u, u)}$  des inerties  $I_{\Delta_u^\perp}$  associées à  $q$  directions

$M$ -orthogonales  $\Delta_u$  de l'espace euclidien  $(E, M)$  des individus. Ils sont immédiatement transposables en substituant à  $(E, M)$  l'espace  $(F, D_p)$  des variables ; on est alors amené à considérer les quotients de Rayleigh  $\frac{D_p(W D_p c, c)}{D_p(c, c)}$ . Remarquons qu'alors les domaines des systèmes de quo-

tients de Rayleigh associés à des directions  $M$ -orthogonales dans  $E$  coïncident avec ceux associés à des variables  $D_p$ -orthogonales dans  $F$  si celles-ci sont combinaisons linéaires des variables  $x^j$  ; on a alors, si  $c = X'u$  :

$$\frac{D_p(W D_p c, c)}{D_p(c, c)} = \frac{V(MV u, u)}{V(u, u)} .$$

### 2.6 Remarque sur le domaine du spectre des opérateurs AVA'M

Dans un espace  $E$  de métrique  $M$ , les carrés des longueurs des axes de l'ellipsoïde associé à l'ajustement AVA' sous-dominant de  $V$  de rang  $q$ , étudié dans la première partie, sont les  $q$  moments principaux  $\mu_i$  de l'analyse en composantes principales du triplet  $(AX, M, D_p)$  ; on montre sans difficulté que :

$$\mu_i = \frac{V(MV v_i, v_i)}{V(v_i, v_i)}$$

où les vecteurs  $v_i$  sont les images par  $V^{-1}$  des vecteurs propres de AVA'M .

Aussi, quand  $A$  parcourt l'ensemble des projecteurs  $V^{-1}$ -symétriques de rang  $q$ , le domaine  $D_q$  des systèmes de moments principaux  $(\mu_1, \dots, \mu_q)$  est un sous-ensemble du domaine  $C_q$  des systèmes de quotients de Rayleigh. Les théorèmes 4 et 5 de la première partie de cet article permettent de

préciser ce sous-ensemble.

Au sens du  $M$  préordre  $O_2$  (comparaison, dans la métrique  $M$ , des longueurs des axes des ellipsoïdes), la forme  $AVA'$  est inférieure à l'ajustement obtenu pour l'opérateur  $A$  de  $M$ -projection sur le sous-espace engendré par  $q$  premiers axes principaux. Cette forme étant de plus supérieure à l'ajustement correspondant aux  $q$  derniers axes principaux, on a :

$$(4) \quad \lambda_{i+p-q} \leq \mu_{(i)} \leq \lambda_i \quad \text{pour } i=1,2,\dots,q$$

en notant  $\mu_{(i)}$  les moments principaux réordonnés par valeurs décroissantes.

Ces inégalités sont aussi une conséquence du théorème de séparation de STURM. La réciproque de ce théorème [Glazman-Liubitch, 1972 p. 163] montre que les inégalités (4) caractérisent le domaine  $D_q$  des spectres des  $AVA'M$ . La figure 3 illustre dans des cas simples le domaine  $C_q$  des systèmes de quotients de Rayleigh et le domaine  $D_q$  des systèmes de moments principaux  $(\mu_1, \mu_2, \dots, \mu_q)$ .

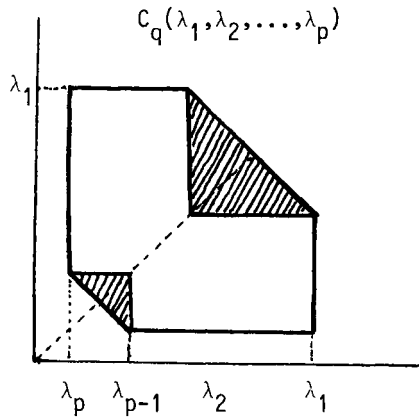


Figure 3

Relation entre les domaines  $D_q$  des systèmes de valeurs propres et  $C_q$  des systèmes de quotients de Rayleigh pour  $q=2$ ,  $p>4$ .  $D_2$  et  $C_2$  coïncident sauf pour la partie grisée (correspondant aux systèmes de quotients de Rayleigh ne pouvant s'interpréter comme systèmes de valeurs propres des  $AVA'M$ ).



### 3. OPTIMALITE ET ANALYSE FACTORIELLE SOUS CONTRAINTES

On rappelle que le modèle factoriel classique admet, sous certaines conditions, une seule infinité de solutions [ANDERSON - 1956], les systèmes de composantes factorielles se déduisant les uns des autres par des rotations dans l'espace des variables. L'introduction des contraintes peut déjà apparaître, dans ce contexte, comme un moyen de lever l'indétermination mathématique inhérente aux spécifications de ce modèle. Mais elles n'ont pas ce seul rôle ; dès que l'on veut aller au-delà des critères purement mathématiques, pour rechercher des dimensions qui réalisant des compromis entre certaines qualités de l'information traitée et ce que connaît par ailleurs l'analyste, on peut avoir recours aux contraintes.

De façon beaucoup plus générale, les contraintes seront utilisées pour introduire dans une méthode d'analyse de données certaines connaissances a priori pour en améliorer les résultats ou même pour mettre à l'épreuve certaines hypothèses ; il s'agira de rendre les résultats plus interprétables, plus opérationnels, c'est-à-dire à la fois plus robustes, plus stables dans le temps et dans l'espace, et surtout plus clairs et plus utiles. On glisse ici d'une optique purement exploratoire à une optique que l'on qualifie de "confirmatoire" ; on passe de l'inférence statistique à la démarche scientifique classique qui repose sur l'expérience : les résultats obtenus viennent ou non confirmer les a priori introduits dans la procédure même de l'analyse [GUTTMAN - 1985].

Les psychométriciens ont été les premiers à donner une base scientifique à cette démarche en introduisant la notion de "structure simple" [THURSTONE - 1935] et en proposant des techniques pour en obtenir : il s'agissait de mettre en évidence des composantes factorielles telles que la matrice de leurs corrélations avec les variables observées soit "simple", c'est-à-dire constituée avec un maximum de zéros ou de valeurs proches de 1 . Notons que ces exigences, ou contraintes, se traduisent par l'appartenance des composantes factorielles à des sous-espaces (contraintes linéaires) ou des cônes (contraintes non linéaires de voisinage) de l'espace des variables. En analyse en Composantes

Principales, où le système de composantes factorielles est en général unique, l'introduction de contraintes obligera le praticien, nous le verrons, à faire un choix entre l'orthogonalité des axes dans l'espace des individus  $E$  (métrique  $M$ ) et la non corrélation des variables dans  $F$  (métrique  $D_p$ ) ; c'est à ce prix qu'il pourra trouver un compromis entre les qualités mathématiques et le caractère opérationnel des résultats.

Les éléments introductifs aux analyses en composantes principales sous contraintes (linéaires, voisinage, ordre, ...), que nous présentons dans les paragraphes suivants, s'appuient principalement sur le résultat général d'optimalité des éléments propres d'un opérateur  $M$ -symétrique positif, énoncé dans le théorème 1 du chapitre 2.

### 3.1 Analyses orthogonales et obliques

En analyse en composantes principales, les contraintes peuvent évidemment être spécifiées soit dans l'espace des variables soit dans celui des individus. Il est notamment aisé de montrer [SCHEKTMAN - 1984] que le type des contraintes, est invariant dans le transport par dualité de ces contraintes d'un espace dans l'autre. Pour introduire ces analyses, prenons un exemple dans l'espace des individus.

Sur la figure 4, on a représenté l'ellipse de concentration et les deux axes principaux d'un triplet  $(X, M, D_p)$  faisant intervenir deux variables. Parmi tous les repères orthogonaux ayant un axe inclus dans le cône hachuré, il est immédiat que  $(\Delta v_1, \Delta v_2)$ , où  $\Delta v_1$  est confondu avec la frontière du cône la plus proche de la deuxième direction principale  $\Delta u_2$ , est le repère, respectant la contrainte, le moins éloigné de celui des axes principaux. Notons qu'ici  $\Delta v_1$ , l'axe principal respectant la contrainte, n'est pas celui qui a l'inertie maximum.

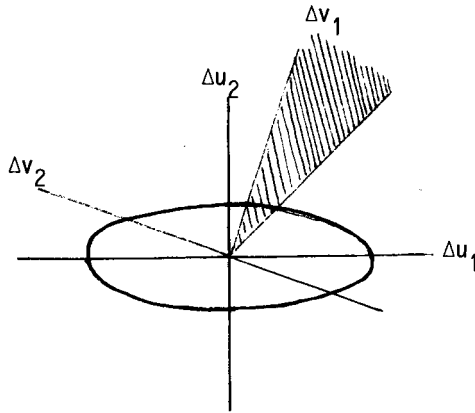


Figure 4

Axes sous une contrainte de voisinage

Cet exemple montre que la démarche pas à pas de détermination des axes principaux n'est pas adaptée à la résolution de contraintes : au lieu de les considérer séparément, il apparaît au contraire nécessaire de les considérer dans leur ensemble, c'est-à-dire de caractériser globalement le repère des axes principaux. Pour atteindre l'objectif on aura par exemple recours à un critère directionnel [tableau 1 du chapitre 2] : dans l'ensemble supra le critère "rendre  $I_{\Delta v_1}^2 + I_{\Delta v_2}^2$  maximum",

qui respecte la S-convexité stricte, conduit au couple  $(\Delta v_1, \Delta v_2)$  indiqué sur la figure 4.

Avant d'aller plus loin rappelons brièvement quelques résultats sur les orthogonalités des directions (axes ou composantes) principales. On sait que les axes principaux du triplet  $(X, M, D_p)$  sont  $M$  et  $V^{-1}$  orthogonaux ; en fait, ils sont orthogonaux relativement à tous les produits scalaires  $T$  rendant  $TVM$  symétrique. Ainsi, si  $P_j$  est l'opérateur de projection orthogonale sur la  $j^{\text{ième}}$  direction propre de  $VM$  (supposé régulier), alors les axes principaux sont orthogonaux relativement à tous les produits scalaires de la forme  $M \Sigma \alpha_j P_j$  avec  $\alpha_j > 0$ . On a la pro-

priété analogue pour les composantes principales. Il est bien connu que la double orthogonalité, relativement à  $M$  et  $V^{-1}$ , est une propriété caractéristique des axes principaux ; en termes équivalents, si une base  $M$ -normée de vecteurs  $u$  de l'espace  $E$  est  $M$ -orthogonale et si les variables  $c = X'Mu$ , associées dans  $F$  par dualité, sont  $D_p$ -orthogonales alors les vecteurs  $u$  engendrent un système d'axes principaux de  $(X, M, D_p)$ , et les vecteurs  $c$  un système de composantes principales. En fait, il est facile de montrer que les directions principales sont caractérisées par la double orthogonalité relativement à une infinité de couples de produits scalaires ; c'est notamment le cas pour les couples de produits scalaires de la forme  $M \sum \alpha_j P_j$  et  $D_p \sum \beta_k Q_k$  avec  $\beta_k > 0$ ,  $\alpha_j \beta_j = 1/\lambda_j$ , où  $Q_k$  est l'opérateur de projection orthogonale sur la droite engendrée par la  $k$ ème composante principale.

Dès que l'on impose des contraintes les directions que l'on cherche ne peuvent plus satisfaire, sauf exception (directions principales respectant les contraintes), toutes ces conditions d'orthogonalité et en particulier simultanément celles qui impliquent  $M$  et  $D_p$ .

Définition 4 : Analyses orthogonales et obliques

On dit qu'une A.C.P. de  $(X, M, D_p)$  sous contraintes est

- orthogonale, si les composantes principales sont supposées non corrélées ( $D_p$ -orthogonalité) ;
- oblique, si les axes principaux sont supposés  $M$ -orthogonaux.

Cette terminologie est cohérente avec celle utilisée en analyse factorielle classique pour définir les structures simples ou introduire les algorithmes de rotation orthogonale ou de rotation oblique. Rappelons que ces derniers, proposés ultérieurement, sont bien adaptés à la recherche d'une typologie de variables. Dans le cas de composantes factorielles obliques l'originalité de la définition proposée réside dans

l'hypothèse d'orthogonalité des axes principaux au sens d'une métrique que l'on se donne a priori et non pas que l'on découvre a posteriori.

### 3.2 Quelques analyses utiles - Optimalité sous contraintes

Introduisons sous forme de propositions les quotients de Rayleigh, dont certains sont bien connus [CAILLIEZ et PAGES - 1976], permettant de mener à bien des analyses orthogonales ou obliques.

PROPOSITION 2 (quotients de Rayleigh pour les analyses obliques) :

Si  $(u_1, \dots, u_p)$  est une base orthogonale de l'espace des individus, alors :

$$I_{\Delta u_j}^{\perp} = \frac{M(VMu_j, u_j)}{M(u_j, u_j)} = \text{var}(c^j)$$

où :  $c^j = X'Mu_j$  .

Pour les analyses obliques les inerties  $(I_{\Delta u_1}^{\perp}, \dots, I_{\Delta u_p}^{\perp})$  constituent donc un p-uple de quotients de Rayleigh associés à un système de p axes M-orthogonaux. De façon plus précise [SCHEKTMAN - 1978] :

PROPOSITION 3 :

Si  $(u_1, \dots, u_p)$  est une base M-normée de l'espace des individus alors les deux assertions suivantes sont équivalentes :

- .  $(u_1, \dots, u_p)$  est M-orthogonale ,
- .  $\text{Cov}[c^j, c^k] = M(VMu_j, u_k)$  .

Si dans les analyses obliques (M-orthogonalité) les critères basés sur les indices de variance et d'inertie sont équivalents, dans les analyses orthogonales ( $D_p$ -orthogonalité) cette propriété n'est généralement

pas vrai : on peut montrer [SCHEKTMAN - 1978] que l'on a toujours :  
 $I_{\Delta^{\perp}u_j} > \text{Var}[c^j]$  . Les points de vue de PEARSON dans E et de

HOTELLING dans F sont alors différents et conduisent à des familles distinctes d'analyses orthogonales. Pour illustrer ces propos quelques analyses utiles sont proposées dans le tableau 2 . A l'exception du critère d'inertie pour les analyses orthogonales, la justification de ces critères découle immédiatement du théorème 1 du chapitre 2, de la proposition 2 et de la proposition 4 suivante qui fournit deux types de quotients de Rayleigh utiles dans les analyses orthogonales.

PROPOSITION 4 : (quotients de Rayleigh pour les analyses orthogonales):

Si les vecteurs forment une base M-normée de l'espace des individus et si les vecteurs  $c$ , où  $c = X'Nu$  forment une base orthogonale, alors les indices  $\text{var}(c)$  et  $I_{\Delta^{\perp}u} \times \text{var}(c)$  forment des p-uples de quotients de Rayleigh, respectivement pour les opérateurs  $WD_p$  et  $(WD_p)^2$  .

Plus précisément :

$$\text{var}(c^j) = \frac{D_p(WD_p c^j, c^j)}{D_p(c^j, c^j)} \quad ,$$

$$I_{\Delta^{\perp}u_j} \times \text{var}(c^j) = \frac{D_p[(WD_p)^2 c^j, c^j]}{D_p(c^j, c^j)} \quad .$$

La démonstration s'effectue à partir de l'expression :  $u = \frac{X D_p c}{D_p(c, c)}$  ,

qui découle de l'orthogonalité des variables  $c$  .

Dans le tableau 2 les liens avec les résultats synthétiques du tableau 1, que l'on trouve au chapitre 2, sont indiqués en commentaires.

	ANALYSES			
	OBLIQUES	ORTHOGONALES		
	(Inertie=Variance)	Inertie	Variance	Mixte
CRITERES	$\sum_{j=1}^q (I_{\Delta u_j}^\perp)^2$	$\sum_{j=1}^p I_{\Delta u_j}^\perp$	$\sum_{j=1}^q (\text{Var}[c^j])^2$	$\sum_{j=1}^q (I_{\Delta u_j}^\perp \text{Var}[c^j])^2$
COMMENTAIRES	DP(1,...,q) MAX (S-convexité stricte)	DG MIN (voir SCHEKTMAN1978)	DP(1,...,q)MAX (S-convexité stricte)	DP(1,...,q) MAX (S-convexité stricte)

Tableau 2 : critères utiles pour les analyses sous contraintes.

3.3 En guise de conclusion : un exemple de problème d'optimisation sous contraintes.

Si on se place dans le cas des analyses obliques voici un exemple de recherche d'axes principaux sous contrainte :

$$\text{maximiser } \sum_{j=1}^m (I_{\Delta v_j}^\perp)^2$$

. avec :  $M(v_j, v_k) = \delta_{jk}$  ;

. sous la contrainte :  $\cos(v_\ell, y) > 0,3$  ;

où  $y$  est une variable donnée et  $\ell$  une valeur entière prise entre 1 et  $m$ .

Pour ce type de contraintes non linéaires aucune solution analytique exacte n'a été proposée à notre connaissance ; seules des solutions algorithmiques approximatives ont été proposées [SCHEKTMAN - 1976, CROQUETTE - 1980 , IBRAHIM - 1982].

Par contre on ne rencontre pas ces difficultés dans la plupart des cas pour les contraintes linéaires où l'on est conduit presque toujours à diagonaliser des opérateurs [CARLIER et al. - 1984]. Rappelons entre autres que l'A.C.P. sous contrainte de non corrélation et l'A.C.P. sur variables instrumentales [RAO - 1964, BONIFACE, ESCOUFIER et al. - 1984, SABATIER - 1984] entrent dans cette dernière catégorie.



## BIBLIOGRAPHIE

- [ 1 ] ANDERSON T.W. (1956). Statistical inference in factor analysis. Proceedings of the third Berkeley Symposium in Mathematical Statistics and Probability.
- [ 2 ] BONIFACE L., ESCOUFIER Y., GONZALEZ P.L., SABATIER R. (1984) : Choix de variables en analyse en composantes principales. R.S.A., 32, 2, 2-15.
- [ 3 ] CAILLIEZ, F. et PAGES J.P. (1976) : Introduction à l'Analyse des Données, SMASH-BURO-ASU, Paris.
- [ 4 ] CARLIER A., CROQUETTE A., SCHEKTMAN Y. (1984) : Analyses en composantes principales sous contraintes linéaires. Rapport technique 8302, E.N.S.A.M.-I.N.R.A.-U.S.T.L., Unité de Biométrie, Montpellier.
- [ 5 ] CROQUETTE A. (1980) : Quelques résultats synthétiques en analyse des données multidimensionnelles : Optimalité et métriques à effets relationnels. Thèse 3ème Cycle, Université Paul Sabatier, Toulouse.
- [ 6 ] CROQUETTE A., KOBILINSKY A., PAGES J.P., SCHEKTMAN Y. (1984) : Optimalité en analyse factorielle : Décomposition de la matrice de variance. Préordre entre formes quadratiques et optimalité. Cahiers du BUR0 n° 43, Université Pierre et Marie Curie, Paris.
- [ 7 ] GLAZHIAN I., LIUBITCH Y. (1972) : Analyse linéaire dans les espaces de dimensions finies. Ed. MIR, Moscou.
- [ 8 ] GUTTMAN L. (1985) : What is not what in statistics. B.M.S., n° 4, 3-28.
- [ 9 ] HOTELLING H. (1933) : Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24, 417-441 et 498-520.
- [10] IBRAHIM A. (1982) : Quelques applications des analyses en composantes principales sous contraintes et des métriques à effets relationnels en analyse des données. Thèse 3ème Cycle, Université Paul Sabatier, Toulouse.
- [11] KOBILINSKY A. (1979) : Décomposition de formes quadratiques en analyse des données. Thèse 3ème Cycle, Université de Paris Sud, Orsay.
- [12] KOBILINSKY A. (1979) : Ordre entre formes quadratiques, application à l'optimalité de sous-espaces en analyse des données. R.S.A., Vol. XXVII, 1, 45-54.

- [13] OSTROWSKI A. (1952) : Sur quelques applications des fonctions convexes et concaves au sens de I. SCHUR. J. Math. pures et appliquées 31, 253-292.
- [14] PEARSON K. (1901) : On lines and planes of closest fit to system of points in space. Phil. Mag., ser. 2, 6, 559-572.
- [15] RAO C.R. (1964) : The use and interpretation of principal component analysis in applied research. Sankhya, Ser. A, 26, 329-359.
- [16] SABATIER R. (1984) : Quelques généralisations de l'ACPVI, Stat. et An. Don., 9, 3, 75-103.
- [17] SCHEKTMAN Y. (1976) Rapport A.T.P.-CNRS, 729910, Paris.
- [18] SCHEKTMAN Y. (1978) : Contribution à la mesure en facteurs dans les sciences expérimentales et à la mise en oeuvre automatique des calculs statistiques. Thèse d'Etat, Université Paul Sabatier, Toulouse.
- [19] SCHEKTMAN Y., HAIT J.R., IBRAHIM A. (1984) : Analyses en composantes principales sous contraintes. Applications. Data Analysis and Informatics, III, 85-97 (E. DIDAY et al. ed.) North Holland, Amsterdam .
- [20] THURSTONE L.L. (1935) : The vectors of the mind. University Chicago Press, Chicago.