

Record Number: 1150

Author, Monographic: Chanut, J. P.//Bertrand, R.//Bobée, B.//Côté, R.//Lachance, M.
PetitJean-Roger, T.

Author Role:

Title, Monographic: La méthode STATIS. Tome 1 : aspects théoriques

Translated Title:

Reprint Status:

Edition:

Author, Subsidiary:

Author Role:

Place of Publication: Québec

Publisher Name: INRS-Eau

Date of Publication: 1980

Original Publication Date: Janvier 1980

Volume Identification:

Extent of Work: vi, 61

Packaging Method: pages et un annexe

Series Editor:

Series Editor Role:

Series Title: INRS-Eau, Rapport de recherche

Series Volume ID: 114

Location/URL:

ISBN: 2-89146-111-8

Notes: Rapport annuel 1979-1980

Abstract: Rapport rédigé en collaboration avec les membres du groupe Méthodes
statistiques de l'INRS
10.00\$

Call Number: R000114

Keywords: rapport/ ok/ dl

LA METHODE STATIS: Tome 1
Aspects théoriques

RAPPORT SCIENTIFIQUE No 114: Tome 1

présenté par

Jean-Pierre Chanut*
(UQAR)

en collaboration avec les membres du groupe
METHODES STATISTIQUES DE L'INRS

INRS-Eau
Université du Québec
C.P. 7500, Sainte-Foy
Québec
G1V 4C7

Janvier 1980

* Représentant l'INRS-Océanologie au sein du
Groupe Méthodes Statistiques de l'INRS

GROUPE METHODES STATISTIQUES DE L'INRS:

Richard Bertrand, INRS-Education

Rudolf Bertrand, INRS-Pétrole

Bernard Bobée, INRS-Eau

Jean-Pierre Chanut, UQAR

Robert Côté, Université Laval

Marius Lachance, INRS-Eau

Thierry PetitJean-Roget, INRS-Santé

* A paraître: LACHANCE, M. et T. PETITJEAN-ROGET (1980).

La méthode STATIS: Tome 2. Programme informatique et exemple d'application. INRS-Eau, rapport scientifique No 115. Groupe méthodes statistiques, 128 p.

AVANT-PROPOS

Ce texte s'appuie en grande partie sur la thèse de L'Hermier des Plantes (1976) qui traite de l'analyse des tableaux à 3 indices. Ce travail qui forme la base du présent rapport est cependant présenté dans une perspective plus large et sous une forme aisément accessible pour l'utilisateur.

De plus, un exemple simplifié est traité en détail au tome 2 pour bien mettre en évidence les diverses étapes de la méthode STATIS ("Structuration des tableaux à trois indices de la Statistique").

Une notation plus homogène et plus explicite que dans la thèse de L'Hermier des Plantes est proposée dans le présent document; cette notation est cohérente avec celle proposée dans l'ouvrage fondamental de Cailliez et Pagès (1976): "Introduction à l'analyse des données".

Pour faciliter une meilleure compréhension, les différents schémas de dualité associés aux différentes étapes de la méthode STATIS sont placés en annexe.

On se référera à la thèse de L'Hermier (1976) pour les démonstrations de théorèmes propres à la méthode STATIS et à l'ouvrage de Cailliez et Pagès (1976) pour les résultats plus généraux propres à l'analyse linéaire.

La méthode est présentée par une optique descriptive seulement, faisant essentiellement appel à la technique de l'Analyse en Composantes Principales (ACP).

La présentation de la méthode dans une optique de prévision, faisant alors appel aux techniques de l'Analyse Canonique, de l'Analyse Factorielle Discriminante et de l'Analyse Factorielle des Correspondances, fera l'objet de travaux ultérieurs. Il en sera de même pour le traitement des données complexes (au sens mathématique du terme) non inclu dans le présent document.

Ce travail a pu être accompli grâce à une subvention de l'INRS (Fonds du président, CRSNG).

La méthode STATIS: Aspects théoriques

	Page
I. INTRODUCTION	
1. Objectif de la méthode	1
2. Présentation générale de STATIS	3
3. Schéma de la méthode	4
4. Obtention des structures	5
II. LES TABLEAUX DE DONNEES	
1. Proximités originales	6
2. Proximités calculées	8
III. LE CUBE STANDARD	
1. Définition	12
2. Transformations	13
IV. L'INTER-STRUCTURE	
1. Définitions	19
2. Matrice de proximités entre juges	20
3. Représentation canonique des juges	25
4. Obtention de l'inter-structure	27
5. Référentiels	29

	Page
V. LA K^{IEME} INTRA-STRUCTURE	
1. Représentation canonique des sujets	32
2. Obtention de la k^e intra-structure	33
VI. QUELQUES CONSIDERATIONS SUR LES STRUCTURES	
1. Structures et opérateurs	35
2. Structures et normalisation	37
VII. L'INTRA-STRUCTURE DE REFERENCE (COMPROMIS)	
1. Définitions de la matrice de référence	41
2. Propriétés de la matrice de référence	42
3. Obtention de l'intra-structure de référence	43
VIII. L'ANALYSE DES DIFFERENCES	
1. Calcul des différences entre sujets	44
2. Visualisation des différences	49
3. Mesure de la qualité de la représentation	50

	Page
IX. L'ANALYSE DES EVOLUTIONS	
1. Détermination des intra-structures dans le référentiel (compromis)	53
2. Visualisation de l'évolution des intra-structures	54
3. Mesure de la qualité des représentations	55
X. EXTENSION DE LA METHODE	
1. Autres matrices de référence	55
2. Autres structures	57
3. Tableaux à plus de trois (3) indices	57
REFERENCES	59
ANNEXE A	

La méthode STATIS

Structuration des Tableaux A Trois Indices de la Statistique

I. INTRODUCTION

1. Objectif de la méthode

La méthode STATIS - Structuration des Tableaux A Trois Indices de la Statistique - permet d'analyser des ensembles de données numériques dits "à trois indices". Dans ces tableaux, un premier indice permet par exemple d'identifier les individus qui font l'objet de l'étude, un second indice repère les variables qui ont été observées sur ces individus tandis que le troisième indice permet d'accéder à différentes situations (ou différents instants) dans lesquelles ces variables ont été observées sur ces individus.

Une autre terminologie, marquée par le développement incessant des méthodes statistiques multidimensionnelles dans les sciences humaines, est également employée pour désigner la nature de chacun des trois indices; à savoir,

- les individus sont encore appelés sujets, les variables sont aussi des caractères observés chez les sujets et les situations correspondent à des juges qui ont observé les différents caractères chez les sujets en question.

Chacun des tableaux "individus x variables" (ou "sujets x caractères") peut faire l'objet d'une analyse en composantes principales qui permettra d'étudier les ressemblances des individus telles qu'elles apparaissent dans ce tableau. La question qui demeure concerne l'évolution des ressemblances des individus lorsqu'on passe d'un tableau à l'autre. La méthode STATIS propose une façon d'aborder la comparaison de plusieurs tableaux de données et définit une vision des individus qui est un bon compromis entre les différentes visions proposées par les différents tableaux.

La particularité de la méthode STATIS consiste à trouver le meilleur compromis entre les différents tableaux correspondant aux différentes situations. En fait, ce compromis apparaissant sous la forme d'une matrice sera non pas une moyenne arithmétique des différentes situations mais une moyenne pondérée, i.e. une combinaison linéaire des différents tableaux. Si toutes les situations conduisent à une vision voisine des individus alors tous les "poids" sont sensiblement égaux et le compromis est proche de la moyenne arithmétique. Si, au contraire, à côté d'un groupe assez homogène de situations, existent une ou deux situations dont les visions sont très particulières, les poids attachés à ces dernières situations vont être très faibles et les tableaux de données correspondants n'interviendront pas dans la définition du compromis. En termes imagés, le compromis recherche ce qui dans les visions est majoritaire; il néglige les minorités éventuelles.

La factorisation canonique ou diagonalisation (recherche des valeurs propres et des vecteurs propres) de la matrice compromis fournit la meilleure représentation des individus qui est encore appelée l'intra-structure de référence dans le cadre de la méthode STATIS. Toutes les études ne justifient pas qu'on étudie cette intra-structure en détail. Il reste dans tous les cas que la projection de toutes les situations peut être faite dans l'espace du compromis (en fait sur les 2 ou 3 premiers axes de cet espace) et qu'il est alors possible d'étudier les différences entre les visions de chacune des situations et celle du compromis ou encore l'évolution des visions d'une situation à l'autre.

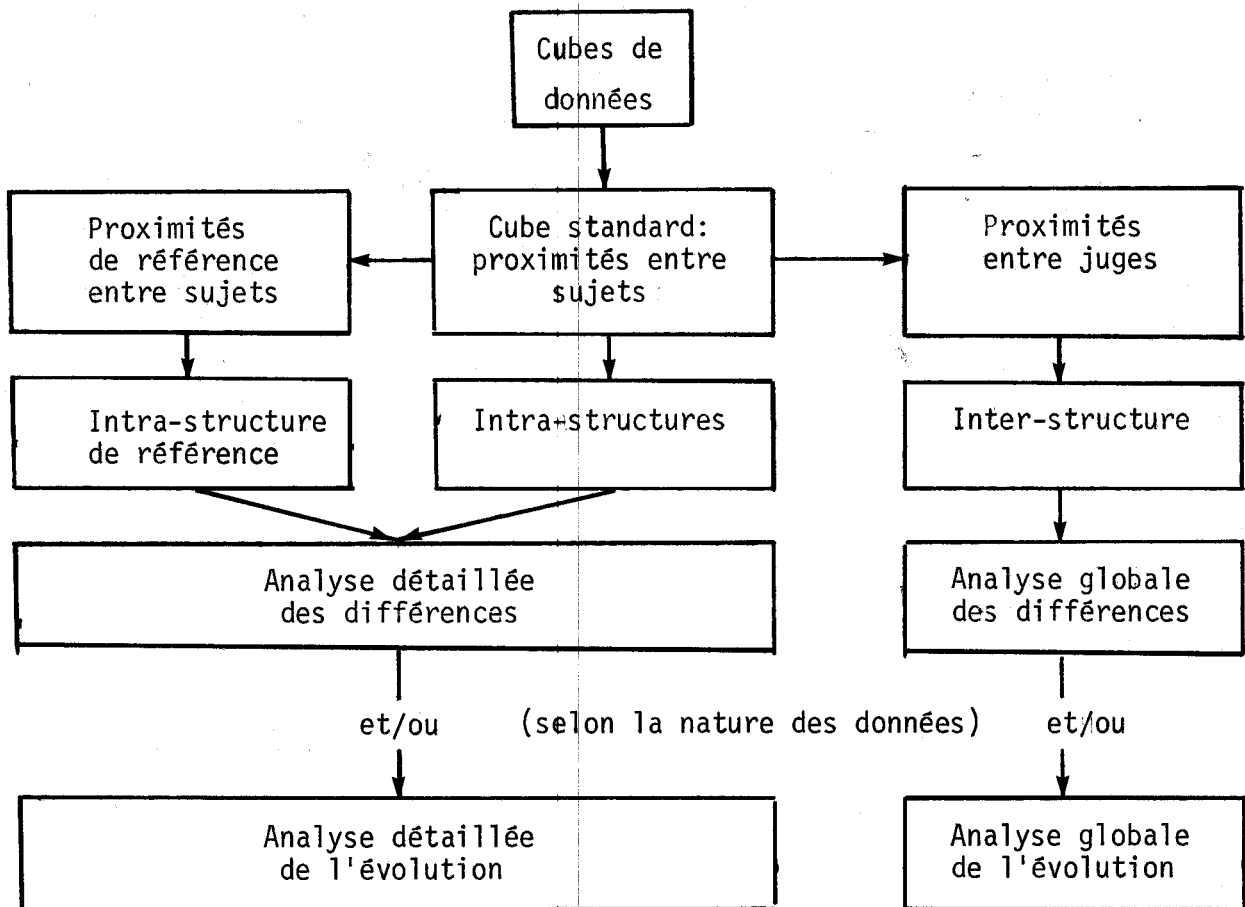
2. Présentation générale de STATIS

La méthode STATIS comprend cinq (5) parties:

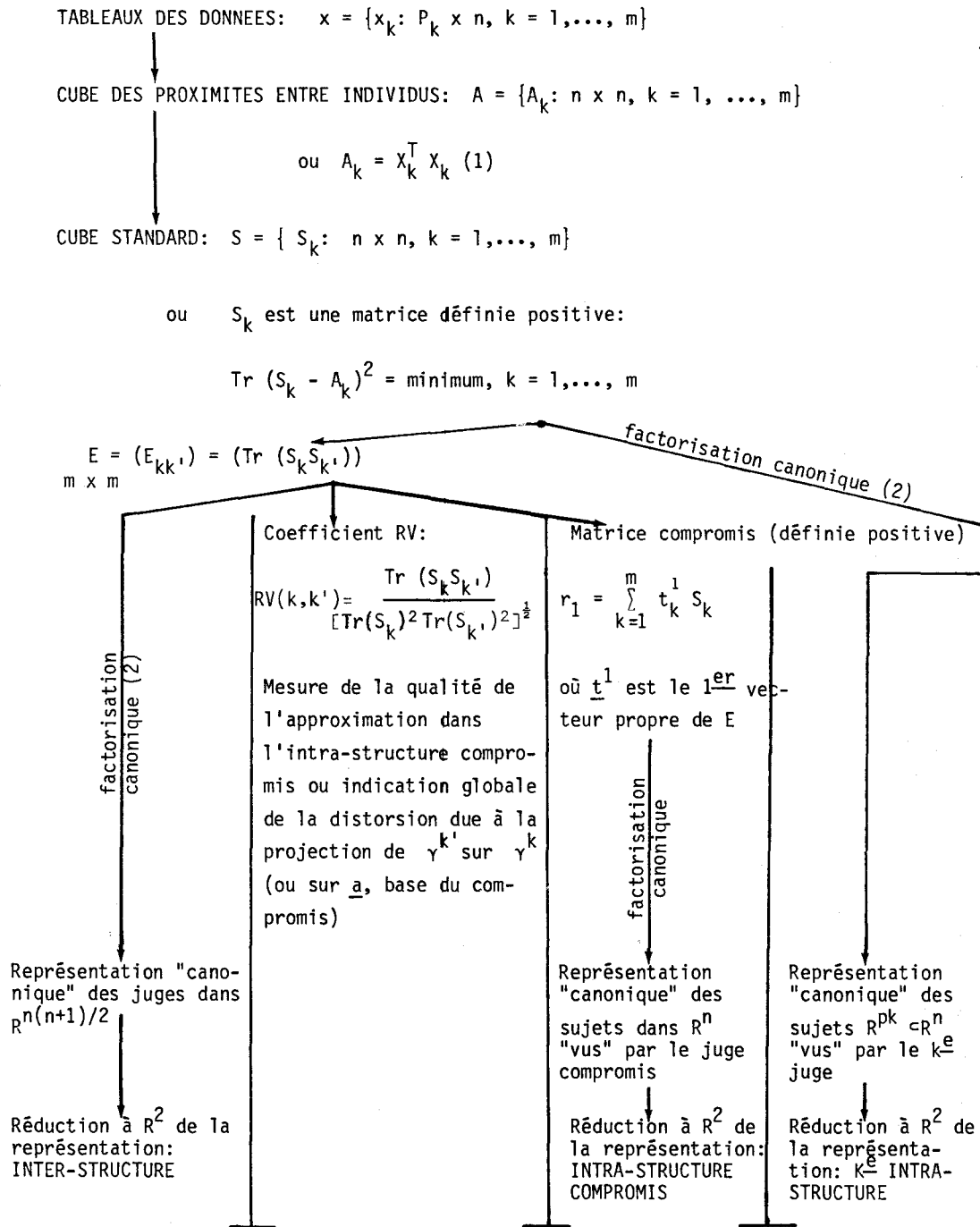
- (1) à partir des données originales $m \times n \times n$ ou $m \times n \times p$, obtention du cube standard $S = \{S_k: n \times n, k = 1, 2, \dots, m\}$ où les S_k seront des matrices définies positives de produits scalaires;
- (2) quantification de la proximité entre les m "juges" et obtention de l'inter-structure;
- (3) obtention de l'intra-structure pour chacun des m "juges" S_k et obtention d'une intra-structure compromis (dite intra-structure de référence);

- (4) analyse des différences entre l'intra-structure pour chaque "juge" et l'intra-structure de référence;
- (5) dans le cas où m est le facteur "temps", on a une série chronologique de tableaux: étude de l'évolution des intra-structures.

3. Schéma de STATIS



4. Obtention des structures: procédure générale



(1) La "transposée" de la matrice Z est désignée par Z^T .

(2) La "factorisation canonique" d'une matrice carrée consiste à "diagonaliser", c'est-à-dire, à rechercher les valeurs propres et vecteurs propres de la matrice. Cette opération constitue une étape importante dans l'analyse en composantes principales.

II. LES TABLEAUX DE DONNEES

Les tableaux ou "cubes" de données à analyser peuvent être de plusieurs types:

C1: similarités	C5: ultramétriques
C2: dissimilarités	C6: ordonnances, notations
C3: distance	C7: profils
C4: produits scalaires	

1. Cubes des proximités originales

Similarités et dissimilarités

Soient I une population de n sujets et K une population de m juges;
 $I = \{ 1, 2, \dots, n \}$ et $K = \{ 1, 2, \dots, m \}$.

Si l'application $g: K \times I \times I \longrightarrow [r_1, r_2] \subset \mathbb{R}$ qui, $\forall k \in K$ et $\forall (i, i') \in I \times I$, fait correspondre le réel $a_{ii'}^k$, de l'intervalle $[r_1, r_2]$ vérifie les deux axiomes:

$$(1) \forall k \in K \text{ et } \forall (i, i') \in I \times I:$$

$$a_{ii'}^k = a_{i'i}^k;$$

$$(2) \left\{ \begin{array}{l} (2a) \forall i \in I \text{ et } \forall k \in K: \\ \quad a_{ii}^k = r_1 \text{ et } a_{ii''}^k < a_{i'i''}^k \iff i \text{ est plus proche de } i'' \text{ que } i'; \\ \quad \text{ou bien} \\ (2b) \forall i \in I \text{ et } \forall k \in K: \\ \quad a_{ii}^k = r_2 \text{ et } a_{ii''}^k < a_{i'i''}^k \iff i \text{ est moins proche de } i'' \text{ que } i'; \end{array} \right.$$

Alors les $a_{ii'}^k$ sont des proximités entre les sujets, et l'ensemble $C. = \{ A_k = (a_{ii'}^k): n \times n, k = 1, \dots, m \}$ est un cube de proximités originales.

Définition 1: L'élément $a_{ii'}^k$ est la proximité donnée par le juge k entre les sujets i et i' .

Définition 2: Si les axiomes (1) et (2a) sont vérifiés, alors les proximités $a_{ii'}^k$ mesurent des dissemblances entre les sujets et sont appelées dissimilarités; $C2 = \{ A_k = (a_{ii'}^k): n \times n, k = 1, \dots, m \}$ où A_k est une matrice de dissimilarités.

Définition 3: Si les axiomes (1) et (2b) sont vérifiés, alors les proximités mesurent des ressemblances entre les sujets et sont appelées similarités; $C1 = \{ A_k = (a_{ii'}^k): n \times n; k = 1, \dots, m \}$ où A_k est une matrice de similarités.

Exemple. Cinq (5) personnes donnent chacune des notes de ressemblance (ou de distance) à dix (10) hommes politiques. Les notes sont comprises entre 0 et 100. On a donc $m = 5$ juges et $n = 10$ sujets; on obtient 5 matrices 10×10 .

- si 0 indique la ressemblance maximum et 100 la ressemblance minimum, alors on obtient des matrices de dissimilarité;
- si 0 indique la ressemblance minimum et 100 la ressemblance maximum, alors on a des matrices de similarité.

Distances. La distance entre sujets est une fonction $d: I \times I \longrightarrow R$ telle que:

- (a) $\forall (i, i') \in (I \times I), d(i, i') \geq 0$
- (b) $d(i, i') = 0$ si $i = i'$
- (c) $\forall (i, i') \in (I \times I), d(i, i') = d(i', i)$
- (d) $\forall i \in I, i' \in I, \forall i'' \in I, d(i, i') \leq d(i, i'') + d(i'', i')$

Ainsi, une distance est une dissimilarité au sens défini précédemment

(avec la propriété (d) en plus). Ex.: $\forall (i, i') \in I \times I: i = i' \iff d(i, i') = 0$ est une dissimilarité nulle. $C3 = \{A_k: n \times n; k = 1, \dots, m\}$ où $A_k = (d_{ii'}^k)$ est une matrice de distances.

2. Cubes des proximités calculées

Produits scalaires

Rappel de la définition: considérons l'application $g: R^n \times R^n \xrightarrow{g} R$, telle que $(\underline{x}, \underline{y}) \rightarrow g(\underline{x}, \underline{y})$ qui à tout couple de vecteurs $(\underline{x}, \underline{y})$ contenu dans l'espace vectoriel produit $R^n \times R^n$ associé le nombre (scalaire):

$$g(\underline{x}, \underline{y}) = \sum_{i=1}^n x_i y_i$$

alors

$$\langle \underline{x}, \underline{y} \rangle = g(\underline{x}, \underline{y}) = \sum_{i=1}^n x_i y_i \quad [\langle \rangle \text{ produit scalaire}]$$

est le produit scalaire des vecteurs $\underline{x}$ et $\underline{y}$. Il vérifie les propriétés suivantes (bilinearité):

$$(1) \quad g(\underline{x}_1 + \underline{x}_2, \underline{y}) = g(\underline{x}_1, \underline{y}) + g(\underline{x}_2, \underline{y})$$

$$(2) \quad g(\alpha \underline{x}_1, \underline{y}) = \alpha g(\underline{x}_1, \underline{y})$$

$$(3) \quad g(\underline{x}, \underline{y}_1 + \underline{y}_2) = g(\underline{x}, \underline{y}_1) + g(\underline{x}, \underline{y}_2)$$

$$(4) \quad g(\underline{x}, \alpha \underline{y}) = \alpha g(\underline{x}, \underline{y}).$$

Matrice de produits scalaires. Le produit scalaire sert à définir des distances entre éléments de $\mathbb{R}^n$.

$$d(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\| = [\langle \underline{x} - \underline{y}, \underline{x} - \underline{y} \rangle]^{\frac{1}{2}} = \left[\sum_{i=1}^n (x_i - y_i)^2 \right]^{\frac{1}{2}}$$

Etant donné qu'une telle distance est une dissimilarité au sens défini plus haut, les m matrices $A_k: n \times n$ peuvent être des matrices de produits scalaires.

Exemple: matrice de covariance.

Ultramétriques (proximités...)

Une ultramétrique (pour le k^{e} juge) entre les sujets est une dissimilarité a_{ij}^k , (pour le k^{e} juge) telle que:

$$a_{ij}^k \leq \max (a_{ij''}^k, a_{i'i''}^k) , \forall i, i', i'' \in I.$$

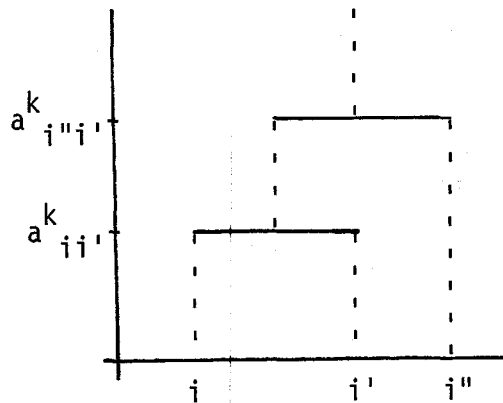
$C_5 = \{A_k: n \times n, k = 1, \dots, m\}$ où A_k est la matrice des proximités ou distances ultramétriques entre les sujets pour le k^e juge.

Exemple: Soient K l'ensemble des méthodes de classification hiérarchique et I un ensemble de n sujets classés par chacune des méthodes.

Considérons pour la k^e méthode, la matrice des distances ultramétriques entre les sujets est $A_k = (a_{ii'}^k): n \times n$. Ses éléments vérifient:

$$(1): a_{ii'}^k = a_{i'i}^k$$

$$(2a) \left\{ \begin{array}{l} a_{ii}^k = r_1 = 0 \\ a_{ii'}^k < a_{i''i}^k \iff i \text{ est plus proche de } i' \text{ que } i'' \text{ (équivalent à une dissimilarité)}, \text{ ce qui s'assure par simple lecture du } k^e \text{ arbre du dendrogramme:} \end{array} \right.$$



Ordonnances, notations

On a des ordonnances ou des notations quand chacun des m juges ordonne les n sujets. On obtient alors m vecteurs $\underline{n}_k: 1 \times n$ ($k=1, \dots, m$) d'éléments n_k^i correspondant à la note ou au rang du i^{e} sujet attribué par le k^{e} juge; le cube de notations est l'ensemble

$$C6 = \left\{ N_k: 1 \times n; k = 1, \dots, m \right\}.$$

Exemples

1. Chacun des m dégustateurs ordonne n marques de vin.
2. Chacun des m examens ordonne n étudiants.

(voir également L'Hermier 1976, p. 21-22).

Profils

Soient les ensembles I contenant n sujets, J contenant p_k caractères ou variables et K contenant m juges.

Soit $X_k: p_k \times n$ la matrice des mesures x_{ji}^k du j^{e} caractère faite par le juge k sur le sujet i . L'ensemble

$$C7 = \left\{ X_k: p_k \times n; k = 1, \dots, m \right\}$$

est appelé cube de profils.

Exemple. On mesure les p paramètres physico-chimiques de chacun des n échantillons (sujets), et cela est répété pendant m jours consécutifs.

Remarque. L'ensemble des ordonnances ou notations de n sujets par m juges

$$C6 = \{ N^k \in R^n; k = 1, \dots, m \} = \{ N_k: 1 \times n; k = 1, \dots, m \}$$

peut être assimilé à un cube de profils

$$C7 = \{ X_k: p_k \times n; k = 1, \dots, m \} \text{ tel que } p_k = 1, \forall k = 1, \dots, m.$$

III. LE CUBE STANDARD

1. Définition. Le cube standard S de l'analyse conjointe (STATIS) est un cube de matrices de produits scalaires:

$$S = \{ A_k: n \times n; k = 1, \dots, m \}$$

Dans la méthode STATIS, chaque matrice doit être factorisée canoniquement (*) et le résultat de cette factorisation, visualisée. Il est donc équi-

(*) La "factorisation canonique" ou "diagonalisation" d'une matrice carrée consiste à trouver les racines caractéristiques du polynôme associé, c'est-à-dire les valeurs propres et vecteurs propres de cette matrice. Cette opération constitue une étape importante dans l'analyse en composantes principales (A.C.P.).

valent de remplacer chaque matrice carrée A_k , semi définie positive, par la matrice carrée S_k , définie positive, la plus proche au sens des moindres carrés généralisés avant la factorisation. On indiquera alors la qualité de l'approximation ainsi réalisée.

Dans ces conditions, le cube standard de l'analyse conjointe (STATIS) est un ensemble de m matrices carrées S_k de produits scalaires, définies positives:

$$S = \{ S_k: n \times n; k = 1, \dots, m \}.$$

2. Obtention du cube standard

Le graphique de la page suivante donne le schéma de l'ensemble des transformations qui permettent, à partir de données cubiques initiales, d'obtenir le cube standard. Dans ce qui suit, on examine successivement la signification des transformations 1 à 10.

1 Ajout à l'ensemble I des sujets d'un élément appelé sujet-origine qui reçoit obligatoirement la même notation de la part de chaque juge de l'ensemble K.

Le cube de notations $C6 = \{ N_k: 1 \times n; k = 1, \dots, m \}$ devient le cube des profils $C7 = \{ X_k: 1 \times (n + 1); k = 1, \dots, m \}$ où

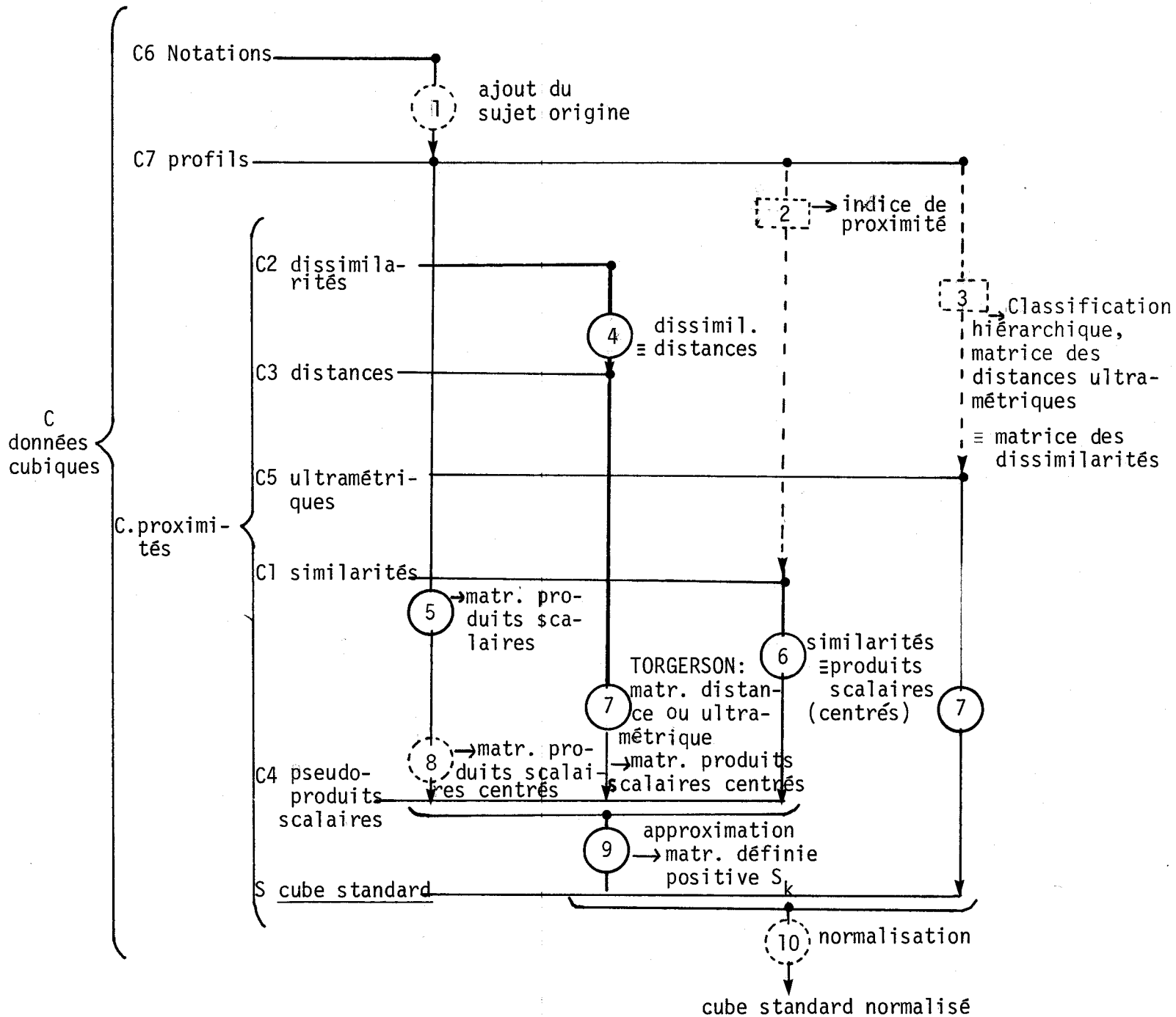
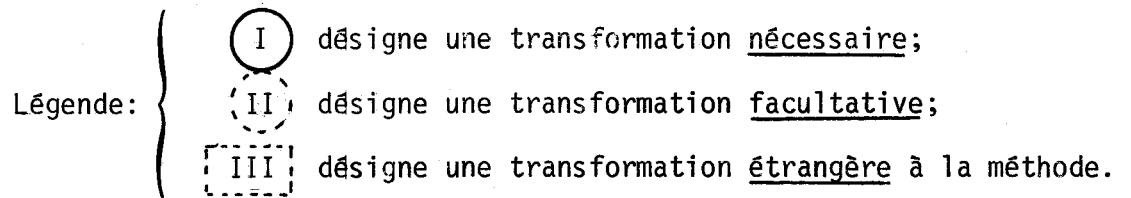
$$x_{1i}^k = n_i^k \text{ pour } i = 1, \dots, n$$

et

$$x_{1i}^k = c: \text{ valeur commune à tout juge } k \text{ pour } i = n + 1 \text{ (sujet-origine).}$$

La justification de cette transformation des données se trouve dans le fait que deux notations N_k et $N_{k'}$, qui ne diffèrent que par une permutation circulaire des composantes, fournissent les mêmes distances entre les sujets.

Schéma de l'ensemble des transformations



De plus, l'interprétation est facilitée par le positionnement de ce sujet-origine dans les intra-structures (Voir exemple de L'Hermier 1976, p. 80-84).

[2] Définition d'un indice de proximité.

Les indices les plus utilisés ont été décrits par Gower (1971), [p. 857-874].

[3] Méthodes de classification hiérarchique fournissant chacune une matrice de distances ultramétriques qui permet habituellement de construire la représentation arborescente.

④ Une dissimilarité expérimentale est assimilée à une distance quelconque:

$$a_{ii'}^k \equiv d_{ii'}^k, \quad \forall i, i' \in I \text{ et } k \in K.$$

⑤ Définition d'un produit scalaire qui permet de transformer un cube de profils en un cube de proximités.

Etant donné un cube de profils $C7 = \{X_k: p_k \times n; k = 1, \dots, m\}$ (éventuellement $p_k = 1, \forall k = 1, \dots, m$) et une forme quadratique $M_k: p_k \times p_k$, on définit une matrice des proximités calculées $A_k = X_k^T M_k X_k$ à partir de chaque matrice X_k et de la métrique correspondante M_k pour obtenir un cube de proximités calculées (produits scalaires):

$$A = \left\{ A_k : n \times n; k = 1, \dots, m \right\}.$$

Remarque. L'application $A_k = X_k^T M_k X_k$ n'est autre que l'application, notée W_k , du schéma de dualité associé à l'ACP de la matrice des données X_k (Voir annexe 3).

⑥ Une similarité expérimentale est assimilée à un produit scalaire quelconque.

⑦ Transformation de Torgerson.

(Cf. Torgerson 1958, p. 257; Rao 1965, p. 259-358; Escoufier 1974, p. 3).

Soit $C3 = \{ D_k : n \times n; k = 1, \dots, m \}$ l'ensemble des matrices de distances entre n sujets représentés dans un espace vectoriel de dimension finie (ou $C5$). L'élément (i, j) de D_k étant d_{ij}^k , la transformation de Torgerson donne l'élément s_{ij}^k d'une matrice S_k de produits scalaires des vecteurs ayant pour origine le centre de gravité des sujets et pour extrémités les points représentant les sujets:

$$s_{ij}^k = \frac{1}{2n^2} \sum_{i'=1}^n \sum_{j'=1}^n \left[\left(d_{ii'}^k \right)^2 + \left(d_{jj'}^k \right)^2 - \left(d_{ij}^k \right)^2 - \left(d_{i'j'}^k \right)^2 \right]$$

Remarque. Les transformations (5), (6) et (7) produisent des matrices qui ont la forme de matrices de covariance. Cela est voulu car l'analyse conjointe (STATIS) consiste essentiellement à effectuer des ACP, i.e. des factorisations canoniques. Il reste à s'assurer que les matrices sur lesquelles on fera des ACP sont définies positives. La transformation 9 est introduite à cette fin.

- 8) Soit $C4 = \{A_k: n \times n; k = 1, \dots, m\}$; chaque matrice de produits scalaires A_k est remplacée par une matrice $B_k: n \times n$ ($k = 1, \dots, m$) de produits scalaires entre les n sujets de façon que les factorisations canoniques des B_k fournissent des représentations dans R^{p_k} qui soient centrées de la même manière que les matrices issues de la transformation de Torgerson.

Pour $k = 1, \dots, m$; $i = 1, \dots, n$ et $\ell = 1, \dots, n$, on calcule

$$b_{i\ell}^k = a_{i\ell}^k - \frac{1}{n} \sum_{i'=1}^n (a_{ii'}^k + a_{i'\ell}^k) + \frac{1}{n^2} \sum_{i'=1}^n \sum_{\ell'=1}^n a_{i'\ell'}^k$$

- 9) Approximation, au sens des moindres carrés généralisés, de chaque matrice A_k du cube de (pseudo) produits scalaires $C4 = \{A_k: n \times n; k = 1, \dots, m\}$ par une matrice définie positive S_k .

La matrice S_k qui minimise $\text{Tr} (S_k - A_k)^2$ est donnée (cf. Neil et coll., 1973, théorème 2, rappelé en annexes de L'Hermier, 1976, p. 92; voir annexe 4 ci-après), par :

$$S_k = \sum_{i=1}^{n' \leq n} \lambda_k^i \underline{y}_i^k \left(\underline{y}_i^k \right)^T; \quad \text{avec } \lambda_k^i > 0$$

où λ_k^i et $\underline{y}_i^k$ sont les n' premiers éléments propres (valeurs propres non nulles λ_k^i et vecteurs propres normalisés $\underline{y}_i^k$) de la matrice A_k tels que

$$\left(\underline{y}_i^k \right)^T \cdot \underline{y}_i^k = 1.$$

La qualité de l'approximation est mesurée par le pourcentage

$$\frac{\text{Tr } (S_k - A_k)^2}{\text{Tr } (A_k)^2}$$

L'approximation est excellente lorsque cette quantité est égale à zéro; cela signifie que la matrice A_k était déjà définie positive.

- (10) Normalisation du cube standard: chaque matrice S_k est normalisée à l'unité en divisant chacun de ses éléments $s_{i\ell}^k$ par:

$$\|S_k\|_{\mathcal{F}(R^n)} = \|\underline{x}_k\|_{R^n}^{(n+1)/2}$$

où le vecteur $\underline{x}_k$ est relié à la matrice S_k par la relation $j_{i\ell} = \frac{i(i-1)}{2} + \ell$; $i = 1, \dots, n$; $\ell \leq i$.

$$\|\underline{x}_k\|_M^2 = \langle \underline{x}_k, \underline{x}_k \rangle_M = \underline{x}_k^T M \underline{x}_k$$

où la matrice est associée à l'application M du schéma de dualité proposé par L'Hermier (1976).

Sachant que: $\underline{x}_k^T M \underline{x}_k = \text{Tr } (S_k^2)$, on en déduit $\|S_k\| = \sqrt{\text{Tr } (S_k^2)}$

et le cube standard normalisé est donné par:

$$SN = \left\{ S_k / \|S_k\| : n \times n; k = 1, \dots, m \right\}.$$

IV. L'INTER-STRUCTURE

1. Définitions

Représentation canonique

Soit X une population statistique de n individus sur lesquels p caractères ont été observés. On appelle représentation canonique (ou triviale) la représentation des individus dans l'espace $E = \mathbb{R}^p$ des individus; à deux éléments (individus, sujets) semblables de X on fait correspondre deux éléments (points) de $\mathbb{R}^p$ qui sont voisins.

Structure

Etant donnée une représentation canonique d'une population statistique X dans $\mathbb{R}^p$, on appelle structure toute réduction à $\mathbb{R}^2$ de cette représentation obtenue par projection orthogonale des points représentatifs des individus sur un plan contenu dans l'espace vectoriel $\mathbb{R}^p$.

Deux structures particulières seront considérées:

. l'inter-structure:

C'est la structure obtenue à partir d'une représentation canonique des juges.

. l'intra-structure:

C'est la structure obtenue à partir d'une représentation canonique des sujets.

2. Calcul de la matrice des proximités entre juges

Soit le cube standard $S = \{S_k : n \times n; k = 1, \dots, m\}$. Si l'on se place dans l'espace $R^n (n+1)/2$ des juges du schéma de dualité proposé par L'Hermier (1976), un juge est représenté par le vecteur colonne x_k , construit à partir de la matrice symétrique S_k par l'application, notée $j_{i\ell}$, et définie par:

$$j_{i\ell} = \frac{i(i-1)}{2} + \ell; \quad i = 1, \dots, n; \quad \ell \leq i$$

et telle que chaque élément

$$x_{j_{i\ell}}^k = s_{i\ell}^k$$

si bien que

$$S_k = (s_{i\ell}^k) = \begin{bmatrix} s_{11}^k & \text{---} & \text{---} \\ s_{21}^k & s_{22}^k & \text{---} \\ s_{31}^k & s_{32}^k & s_{33}^k \\ \text{---} & \text{---} & \text{---} \end{bmatrix} \quad \text{et } x_k = (x_{j_{i\ell}}^k) = \begin{bmatrix} x_{11}^k \\ x_{21}^k \\ x_{22}^k \\ x_{31}^k \\ x_{32}^k \\ x_{33}^k \\ \text{---} \end{bmatrix}$$

où $x_{j_{il}}^k$ est la "mesure résultante" ou "mesure produit" de la j^e variable (caractère) faite sur les individus i et l par le k^e juge.

La proximité entre deux juges X_k et $X_{k'}$, dans $R^{n(n+1)/2}$ est mesurée par:

$$\langle X_k, X_{k'} \rangle_M = X_k^T M X_{k'} = \langle S_k, S_{k'} \rangle_{\text{Trace}} = \text{Tr} (S_k S_{k'})$$

en employant l'opérateur trace comme métrique. La matrice $M: [n(n+1)/2 \times n(n+1)/2]$, diagonale, est définie par

$$m_{j_{il}} = \begin{cases} 2 & \text{si } i < l \\ 1 & \text{si } i = l \end{cases}$$

On en déduit la matrice $E: m \times m$ des proximités entre les m juges:

$$E = (E_{kk'}) = \left(\text{Tr} (S_k S_{k'}) \right)$$

Remarque. La mesure de la proximité entre deux juges peut être précisée en montrant comment le schéma de dualité de L'Hermier (1976) est relié aux m schémas classiques de l'ACP (voir annexe 2).

Nature de la proximité entre juges

Lorsque nous avons recherché une équivalence entre tableaux des données et opérateurs en analyse linéaire, nous avons établi l'équiva-

lence entre un tableau de données (X_k, M_k, D_p) et l'opérateur U_{X_k} qui lui est associé (par définition) et tel que:

$$U_{X_k} \stackrel{\text{déf.}}{=} W_k D_p = X_k^T M_k X_k D_p$$

Si l'on choisit $M_k = I_{p_k}$ et $M_k = D_p = I_n$, la matrice associée à l'opérateur U_{X_k} n'est autre que

$$U_{X_k} = X_k^T \quad X_k = A_k$$

Si l'approximation de A_k par S_k est telle que $\text{Tr} (S_k - A_k)^2 = 0$, il s'ensuit que l'application S_k équivalente à l'application $A_k = X_k^T M_k X_k = W_k: F^* \longrightarrow F = R^n$ est de rang maximum, égal à n [$\text{rang} (S_k) = \text{rang} (A_k) = \dim (F) = n$ ou bien $\dim \text{Ker} (S_k) = \dim \text{Ker} (A_k) = 0$; voir également Cailliez et Pagès 1976, p. 245].

Ainsi, chaque matrice S_k du cube standard peut donc être associée à l'opérateur U_{X_k} , lui-même associé au tableau X_k obtenu par factorisation canonique de S_k .

En employant l'opérateur trace comme métrique, la proximité entre deux opérateurs, U_{X_k} et $U_{X_{k'}}$, est mesurée par

$$\langle U_{X_k}, U_{X_{k'}} \rangle_{\text{Tr}} = \text{Tr} (S_k S_{k'}) = E_{kk'}$$

Ainsi, la matrice $E: m \times m$ des proximités entre les juges dans R^n $(n+1)/2$

est identique à la matrice des proximités entre opérateurs associés aux tableaux X_k obtenus par factorisation canonique des matrices S_k du cube standard.

Remarque. Pour plus de détails, voir également Cailliez et Pagès 1976, p. 478-479; Escoufier 1977, p. 127-128.

Si X_k désigne le tableau $p_k \times n$ obtenu par factorisation canonique de S_k , élément de S , on peut (par définition: cf. Escoufier 1973) associer à deux juges k et k' un coefficient RV, défini par:

$$RV(X_k, X_{k'}) = \frac{\langle U_{X_k}, U_{X_{k'}} \rangle}{\|U_{X_k}\| \|U_{X_{k'}}\|}$$

et correspondant au cosinus de l'angle entre les opérateurs U_{X_k} et $U_{X_{k'}}$.

Si les matrices $S_1, S_2, \dots, S_m$ sont normalisées (normées de façon que $\|S_k\|_{Tr} = \sqrt{\text{Tr}(S_k^2)} = 1, \forall k = 1, \dots, m$, alors la matrice E des proximités entre les juges est une matrice de coefficients RV avec pour éléments:

$$E_{kk'} = \text{Tr}(S_k S_{k'}) = RV(X_k, X_{k'}).$$

Equivalence entre deux juges S_k et $S_{k'}$

a. Deux tableaux X_k et $X_{k'}$ sont dits équivalents lorsque les opérateurs

associés U_{X_k} et $U_{X_{k'}}$ sont à distance nulle dans l'espace R^m des opérateurs (de Hilbert-Schmidt):

$$\|U_{X_k} - U_{X_{k'}}\| = 0 \iff X_k \sim X_{k'}$$

Remarque. $U_X = W_k D_p = X_k^T M_k X_k D_p = X_k^T X_k = S_k: F \longrightarrow F = R^m$, espace des proximités entre juges, lorsque $M_k = I_{p_k}$ et $D_p = I_m$.

b. Si S_k et $S_{k'}$ ont les mêmes éléments propres, alors

$$d^2(U_{X_k}, U_{X_{k'}}) = 0$$

Pour la démonstration de cette proposition, voir Escoufier 1975 [(ISEE), p. 6-7)]

c. $U_{X_k} = U_{X_{k'}} \iff RV(X_k, X_{k'}) = 1$.

La matrice E est définie positive

En effet, si on pose $Q = \sum_{k=1}^m a_k S_k$ pour tout vecteur $\underline{a} = (a_k) > 0 \in R^m$, on a:

$$\|Q\|^2 = \left\langle \sum_{k=1}^m a_k S_k, \sum_{k'=1}^m a_{k'} S_{k'} \right\rangle = \sum_{k=1}^m \sum_{k'=1}^m a_k a_{k'} \langle S_k, S_{k'} \rangle = \text{Tr}$$

$$\text{donc } \|Q\|^2 = \sum_{k=1}^m \sum_{k'=1}^m a_k a_{k'} E_{kk'} = \underline{a}^T E \underline{a} > 0.$$

Remarques. a. La matrice E n'est pas toujours définie positive (des contre-exemples sont signalés). Il reste à montrer sous quelle(s) condition(s) la matrice E est définie positive.

b. La matrice Q ressemble à la matrice $r_1 = \sum_{k=1}^m t_k^1 s_k$ où

$\underline{t}^1 = (t_k^1) > 0$ est le 1^{er} vecteur propre de E , dont tous les éléments sont de même signe.

3. Représentation canonique des juges (dans R^m)

Factorisation canonique de E

Soit $E: m \times m$ la matrice des proximités entre juges; $E = (E_{kk'}) = (\text{Tr}(S_k S_{k'}))$; $E (m \times m)$ est une matrice réelle symétrique définie positive; et soit $\{\tau_k, \underline{t}^k; k = 1, \dots, m\}$ l'ensemble de ses éléments propres tels que $(\underline{t}^k)^T \underline{t}^k = \tau_k$ et $\tau_k > 0, \forall k = 1, \dots, m$.

La matrice réelle $t: m \times m$, ayant pour colonnes les vecteurs $\underline{t}^k$ normés à τ_k , minimise $\text{Tr}(E - t^T t)^2$; elle vérifie donc l'équation

$$E = t t^T$$

$$E = \sum_{k=1}^m \underline{t}^k (\underline{t}^k)^T \text{ lorsque } (\underline{t}^k)^T \underline{t}^k = \tau_k$$

$$E = \sum_{k=1}^m \underline{t}^k (\underline{t}^k)^T = \sum_{k=1}^m \sum_{k'=1}^m \tau_k t_{k'}^k (t_{k'}^k)^T$$

avec $\tau_k > 0$

$$\text{lorsque } (\underline{t}^k)^T \underline{t}^k = 1 \text{ et } T = \begin{bmatrix} \tau_1 & 0 \\ 0 & \tau_m \end{bmatrix}$$

Pour cette démonstration, voir (L'Hermier 1976, [p. 31 et 92]).

Représentation canonique

La matrice E est une matrice de produits scalaires entre m juges (notés X_k ; $k = 1, \dots, m$; ou S_k) représentés dans un espace vectoriel isomorphe à $\mathbb{R}^m$.

L'expression $E = t t^T$ montre que, pour le choix de la métrique identité sur cet espace, $N = D_p = I_m$; la matrice t :

$$t = (\underline{t}^1, \underline{t}^2, \dots, \underline{t}^m) = \begin{pmatrix} t_1^1 & t_1^2 & \dots & t_1^k & \dots & t_1^m \\ t_k^1 & t_k^2 & \dots & t_k^k & \dots & t_k^m \\ t_m^1 & t_m^2 & \dots & t_m^k & \dots & t_m^m \end{pmatrix}$$

représente les coordonnées des m vecteurs lignes $\underline{t}_k$ (images des m juges) dans $\mathbb{R}^m$; la k^{e} ligne de t , d'éléments $(t_k^1, t_k^2, \dots, t_k^m) = \underline{t}_k$, représente les composantes du vecteur $\underline{t}_k$, image du k^{e} juge dans $\mathbb{R}^m \in \mathbb{R}^{n(n+1)/2}$ (espace des juges) tandis que le vecteur colonne $\underline{t}^k$ représente les com-

posantes des m juges sur le k^{e} axe contenu dans R^m (espace des proximités).

La représentation ainsi obtenue n'est définie qu'à un arrangement près des colonnes de t , ce qui correspond à un arrangement dans la numérotation des axes de l'espace R^m .

4. Obtention de l'inter-structure

La visualisation des proximités entre juges au moyen des représentations planes est obtenue en réduisant à R^2 la représentation canonique précédente.

Le problème de la représentation du nuage des juges dans $R^m \subset R^{n(n+1)/2}$ étant analogue à celui de la représentation du nuage des individus (sujets) dans $E = R^p$ ou du nuage des variables (caractères) dans $F = R^n$, sa résolution sera obtenue de la même façon en effectuant une ACP sur les données correspondantes.

Ainsi, à partir de la représentation canonique $t = (\underline{t}^1, \underline{t}^2, \dots, \underline{t}^m)$ avec $\tau_1 \geq \tau_2 \geq \dots \geq \tau_m$ (les vecteurs propres de $E = (E_{kk})$ sont rangés dans l'ordre décroissant des valeurs propres), nous savons que la réduction à R^2 de la configuration dans R^m la plus fidèle quant au respect des proximités entre éléments S_k représentés est la projection sur le premier plan factoriel.

Etant donné que la variance mesure la dispersion du nuage de points,

la qualité de cette représentation réduite est mesurée par le pourcentage de variance expliquée par les deux premiers facteurs à savoir:

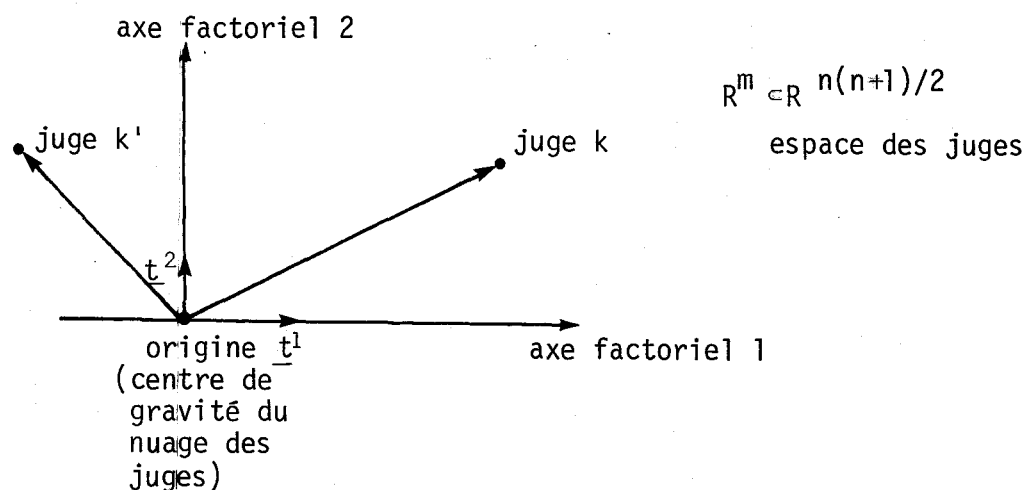
$$(\tau_1 + \tau_2) / \sum_{k=1}^m \tau_k$$

De même, la seconde représentation plane, orthogonale à la première est obtenue en projetant le nuage de points de R^m sur le second plan factoriel défini par les axes principaux 1 et 3 qui expliquent:

$$\frac{(\tau_1 + \tau_3)}{\sum_{k=1}^m \tau_k} \% \text{ de la variance totale}$$

Cet ensemble de projections de la représentation canonique des juges sur les plans factoriels successifs, orthogonaux et ordonnés selon le pourcentage de la variance expliquée visualise bien les proximités entre juges (d'autant mieux que le nombre de plans pris en compte est élevé).
Cet ensemble est appelé l'inter-structure.

Exemple: Visualisation du premier élément $(\underline{t}^1, \underline{t}^2)$ de l'inter-structure définissant le premier plan factoriel dans R^m :



Remarque.

Dans le schéma de dualité proposé par L'Hermier (1976), l'espace des juges est $R^{n(n+1)/2}$, et un juge est représenté par $\underline{x}_k$, la k^{e} colonne de X : $n(n+1)/2 \times m$.

Par contre, dans l'inter-structure proposée après factorisation canonique, les juges sont représentés dans l'espace $R^m \subset R^{n(n+1)/2}$ des juges. Cela tient à l'utilisation de la factorisation canonique qui réduit à m la dimension de l'espace initial $R^{n(n+1)/2}$; de façon générale, $R^m \subset R^{n(n+1)/2}$.

5. Référentiels

Etant donnée la matrice E : $m \times m$ des proximités entre juges associée à l'opérateur WD_p du schéma de dualité proposé par L'Hermier (1976), la diagonalisation de E fournit les éléments propres

$$\{\tau_k, \underline{t}^k; k = 1, \dots, m\}$$

tels que

$$WD_p(\underline{t}^k) = \tau_k \underline{t}^k$$

Définition

On appelle k^{e} référentiel le vecteur de $R^{n(n+1)/2}$ défini par:

$$\underline{r}^k = X.N(\underline{t}^k) = \sum_{k'=1}^m t_{k'}^k \underline{x}_{k'}$$

où $\underline{t}^k \underline{t}^{k'} = \delta_{kk'}$ (Kronecker) et $\underline{x}_{k'}$ est la k^e colonne de la matrice X associée à l'application X du schéma de dualité proposé par L'Hermier (1976).

Remarque. Pour le schéma de dualité associé au k^e vecteur référentiel, voir l'annexe 5.

Propriétés

On montre que:

1. Le vecteur $\underline{r}^k$ est k^e vecteur propre de l'opérateur VM du schéma de dualité proposé par L'Hermier (1976) avec ($k = 1, \dots, m$):

$$VM(\underline{r}^k) = \tau_k \underline{r}^k$$

Démonstration: Cf. L'Hermier (1976), [p. 14].

2. Le vecteur $\underline{r}^k$ maximise $\sum_{k'=1}^m \left[\langle \underline{x}_{k'}, \underline{r}^k \rangle_M \right]^2$ avec ($k = 1, \dots, q$; $q \leq m$), sous la contrainte $(\underline{t}^k)^T \underline{t}^k = \text{constante}$ parmi les vecteurs

$\underline{r}^k$ de la forme: $\sum_{k'=1}^m t_{k'}^k \underline{x}_{k'}$, le maximum étant τ_k^2 pour $\langle (\underline{t}^k)^T, \underline{t}^k \rangle_M = 1$.

Démonstration: Cf. L'Hermier (1976), [p. 15].

3. $\|\underline{r}^k\|_M = \sqrt{\tau_k}$ avec $k = 1, \dots, m$

Démonstration: Cf. L'Hermier (1976), [p. 16].

Remarque. Des propriétés analogues aux précédentes peuvent être déduites à partir des éléments $S_k \in \mathcal{J}(R^n)$ du cube standard S au lieu des éléments $x_k \in R^{n(n+1)/2}$ auxquels ils correspondent dans l'application $j_{i\ell}$: $j_{i\ell} = \frac{i(i-1)}{2} + \ell$; $i = 1, \dots, n$; $\ell \leq i$ (voir ci-dessous, ch. VII, 1) [Cf. L'Hermier (1976), p. 17].

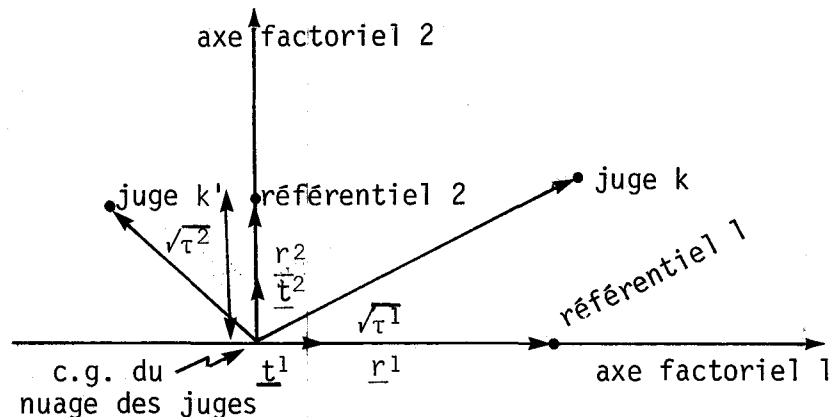
Positionnement des référentiels (dans R^m)

Les vecteurs référentiels $\underline{r}^k$ sont des vecteurs de $R^{n(n+1)/2}$ calculés à partir des vecteurs propres $\underline{t}^k$ par la relation

$$\underline{r}^k = X.N(\underline{t}^k).$$

Par conséquent, le vecteur $\underline{t}^k$ représente le k^e référentiel dans R^n . Etant donné que $\underline{t}^k$ est le k^e vecteur propre de l'opérateur WD_p du schéma de dualité proposé par L'Hermier (1976) avec $(\underline{t}^k)^T \underline{t}^k = 1$, il est aisé de positionner $\underline{r}_k$ dans R^m sur le k^e axe factoriel de l'inter-structure à une distance $\sqrt{\tau_k}$ de l'origine (centre de gravité du nuage des juges).

Exemple: Visualisation des deux premiers référentiels ($\underline{r}^1, \underline{r}^2$) dans le premier plan factoriel de l'inter-structure



$R^m \in R^{n(n+1)/2}$
espace des juges

V. LA $K^{\text{ième}}$ INTRA-STRUCTURE

1. Représentation canonique des sujets (dans R^{p_k})

Factorisation canonique de S_k

Le cube standard S contient, pour chaque juge k ($k = 1, \dots, m$), la matrice carrée $n \times n$: $S_k = X_k^T X_k$ des produits scalaires entre les n sujets.

Soit $\{\lambda_k^j, \underline{y}_j^k; j = 1, \dots, p_k \leq n\}$ l'ensemble des éléments propres de S_k tels que $(\underline{y}_j^k)^T \underline{y}_j^k = \lambda_k^j$.

On sait que [Cf. L'Hermier 1976, p. 92, théorème 2] si A_k est une matrice réelle symétrique et si $\text{Tr} (A_k - S_k)^2$ est minimisée par rapport à une matrice réelle X_k telle que $S_k = X_k^T X_k$, alors

$$S_k = X_k^T X_k$$

$$= \underline{y}^k (\underline{y}^k)^T \quad \text{lorsque } (\underline{y}_j^k)^T \underline{y}_j^k = \lambda_k^j ;$$

$$= \underline{y}^k \Lambda_k (\underline{y}^k)^T = \sum_{\substack{j=1 \\ \lambda_k^j > 0}}^{p_k \leq n} \lambda_k^j \underline{y}_j^k (\underline{y}_j^k)^T$$

$$\text{lorsque } (\underline{y}_j^k)^T \underline{y}_j^k = 1 \quad \text{et} \quad \Lambda_k = \begin{bmatrix} \lambda_k^1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k^{p_k} \end{bmatrix}$$

Représentation canonique

La matrice S_k est une matrice de produits scalaires entre n sujets représentés dans un espace vectoriel isomorphe à R^{p_k} .

L'expression $S_k = \gamma^k \gamma'^k$ signifie que la représentation canonique des sujets pour le k^e juge est définie par:

$${}_{n \times p_k} \gamma^k = (\underline{\gamma}_1^k, \underline{\gamma}_2^k, \dots, \underline{\gamma}_j^k, \dots, \underline{\gamma}_{p_k}^k) = \begin{bmatrix} \gamma_{11}^k & \gamma_{12}^k & \dots & \gamma_{1j}^k & \dots & \gamma_{1p_k}^k \\ \gamma_{i1}^k & \gamma_{i2}^k & \dots & \gamma_{ij}^k & \dots & \gamma_{ip_k}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ \gamma_{n1}^k & \gamma_{n2}^k & \dots & \gamma_{nj}^k & \dots & \gamma_{np_k}^k \end{bmatrix}$$

qui représente les coordonnées des n vecteurs lignes $\underline{\gamma}_i^k$ (images des n sujets) dans R^{p_k} ; la i^e ligne de γ^k , d'éléments $(\gamma_{i1}^k, \gamma_{i2}^k, \dots, \gamma_{ij}^k, \dots, \gamma_{ip_k}^k) = \underline{\gamma}_i^k$, représente les composantes du vecteur $\underline{\gamma}_i^k$, image du i^e sujet dans R^{p_k} , tandis que le vecteur colonne $\underline{\gamma}_j^k$ représente les composantes des n sujets sur le j^e axe contenu dans R^{p_k} .

2. Obtention de la k^e intra-structure

Pour tout juge k ($k = 1, \dots, m$), considérons la représentation canonique des sujets limitée aux p_k ($p_k \leq n$) éléments propres de S_k , de matrice:

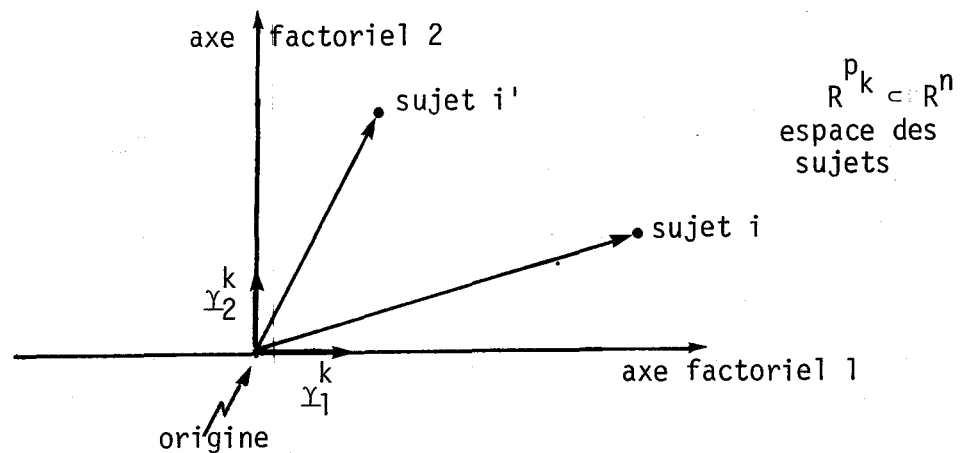
$$Y^k = \begin{pmatrix} Y_1^k & Y_2^k & \dots & Y_j^k & \dots & Y_{p_k}^k \end{pmatrix} = \begin{bmatrix} Y_{11}^k & \dots & Y_{1j}^k & \dots & Y_{1p_k}^k \\ Y_{i1}^k & \dots & Y_{ij}^k & \dots & Y_{ip_k}^k \\ Y_{n1}^k & \dots & Y_{nj}^k & \dots & Y_{np_k}^k \end{bmatrix}$$

$n \times p_k$

et telle que $\lambda_k^1 \geq \lambda_k^2 \geq \dots \geq \lambda_k^j \geq \dots \geq \lambda_k^{p_k} > 0$.

L'ensemble des réductions à R^2 de la configuration des sujets dans R^{p_k} , $\{(Y_1^k, Y_2^k), (Y_1^k, Y_3^k), \dots\}$, ordonnée selon un pourcentage décroissant de la variance expliquée, est appelé la k^e intra-structure. La meilleure représentation plane des sujets vus par le k^e juge au sens de la variance maximum expliquée, est celle dans (Y_1^k, Y_2^k) ; on peut considérer d'autres plans au besoin.

Exemple: Visualisation du premier élément (Y_1^k, Y_2^k) de la k^e intra-structure dans R^{p_k} :



VI. QUELQUES CONSIDERATIONS SUR LES STRUCTURES

1. Structures et opérateurs

L'inter-structure, image de l'espace des opérateurs

Etant donnée l'équivalence entre un tableau de données X_k et l'opérateur U_{X_k} associé, la matrice E $m \times m$, factorisée pour obtenir l'inter-structure, a pour $(i,j)^e$ élément le produit scalaire entre les opérateurs U_{X_k} et $U_{X_{k'}}$ associés respectivement aux tableaux X_k et $X_{k'}$, éléments du cube standard S .

Par conséquent, la représentation canonique des m juges obtenue peut être regardée comme une représentation des m opérateurs U_{X_k} dans un espace contenant l'espace des opérateurs. Dans la base $(\underline{r}^1, \underline{r}^2, \dots, \underline{r}^m)$ définie par les référentiels $\underline{r}^k$ de l'espace représenté (soit $R^{n(n+1)/2}$ pour les éléments X_k , soit $\mathcal{P}(R^n)$ pour les éléments S_k), seul U_{r1} appartient à l'espace des opérateurs positifs U_{X_k} . Cela découle du fait que la matrice X_k , ou plus exactement S_k , a été choisie définie positive. [Cf. L'Hermier 1976, p. 36 et 50 à 52].

De ce fait, par exemple, le cosinus de l'angle de deux vecteurs représentant des opérateurs (correspondant à des juges ou à des référentiels) est égal au coefficient RV entre ces deux opérateurs:

$$\cos \beta_{kk'} = \frac{\langle U_{X_k}, U_{X_{k'}} \rangle_{Tr}}{\|U_{X_k}\|_{Tr} \|U_{X_{k'}}\|_{Tr}} = RV(U_{X_k}, U_{X_{k'}}).$$

Cette propriété reste vraie, que le cube standard soit ou ne soit pas normalisé.

L'intra-structure k , image du spectre de l'opérateur U_{X_k}

La matrice S_k , associée à U_{X_k} , factorisée canoniquement pour obtenir la k^e représentation canonique des sujets, fournit le spectre $\{\lambda_k^j, \gamma_j^k; j = 1, \dots, p_k \leq n\}$ de l'opérateur U_{X_k} .

Les axes $(\gamma_j^k, \gamma_{j'}^k)$ formant le plan (j, j') de la k^e intra-structure contenue dans $R^{p_k} \subset R^n$ servent de repère pour la représentation des sujets dans R^2 . L'expression:

$$(\lambda_k^j + \lambda_k^{j'}) / \sum_{j=1}^{p_k} \lambda_k^j$$

donne le pourcentage de la variance expliquée par ce plan, la variance étant une mesure de la dispersion du nuage des sujets; cette dispersion ne dépend pas du repère, mais seulement des distances entre sujets.

L'ensemble des plans formés par deux axes distincts et orthogonaux $(\gamma_j^k, \gamma_{j'}^k)$ visualise la configuration des sujets dans R^{p_k} , les axes-repères étant les vecteurs propres de S_k . De plus, les plans sont ordonnés selon le pourcentage de variance expliquée de sorte que l'on puisse obtenir une visualisation des sujets aussi précise qu'on le désire avec le moins de plans possible.

Relations entre inter et intra-structures

Deux juges confondus dans l'inter-structure, c'est-à-dire pour lesquels

$d^2(U_{X_k}, U_{X_{k'}}) = 0$, correspondent, d'après l'équivalence 2 (ci-dessus mentionnée, Ch. IV, 2) relative aux matrices S_k , à deux matrices S_k et $S_{k'}$, qui ont mêmes éléments propres, ce qui signifie que les intra-structures k et k' sont identiques. Autrement dit, les distances entre les sujets sont les mêmes, calculées à partir de S_k ou de $S_{k'}$, ou encore, lorsque X_k et $X_{k'}$ sont connues, les analyses en composantes principales de ces tableaux sont équivalentes.

C'est donc en termes de distances entre sujets voisins dans l'intra-structure qu'il faut interpréter la proximité de deux juges dans l'inter-structure.

2. Structures et normalisation

Effet de la normalisation sur l'inter-structure

Nous avons vu (ci-dessus, Ch. IV, 2) que les éléments de la matrice E : $m \times m$ des proximités entre les juges étaient, par définition (Cf. Escoufier 1973):

$$E_{kk'} = \text{Tr}(S_k S_{k'}) = \frac{\langle U_{X_k}, U_{X_{k'}} \rangle_{\text{Tr}}}{\|U_{X_k}\|_{\text{Tr}} \|U_{X_{k'}}\|_{\text{Tr}}} = \text{RV}(X_k, X_{k'}).$$

Lorsque le cube standard S est normalisé, alors:

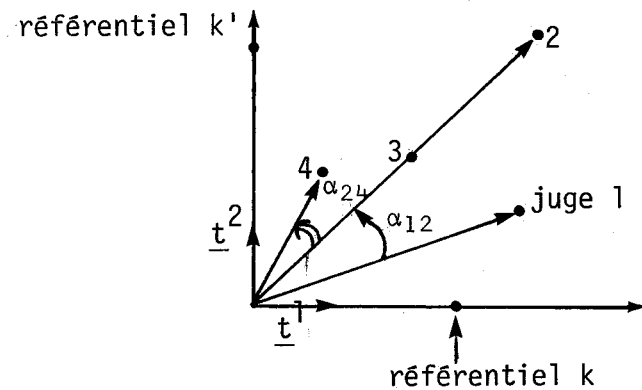
$$\|S_k\|_{\text{Tr}} = \sqrt{\text{Tr}(S_k^2)} = 1, \quad \forall k = 1, \dots, m$$

et

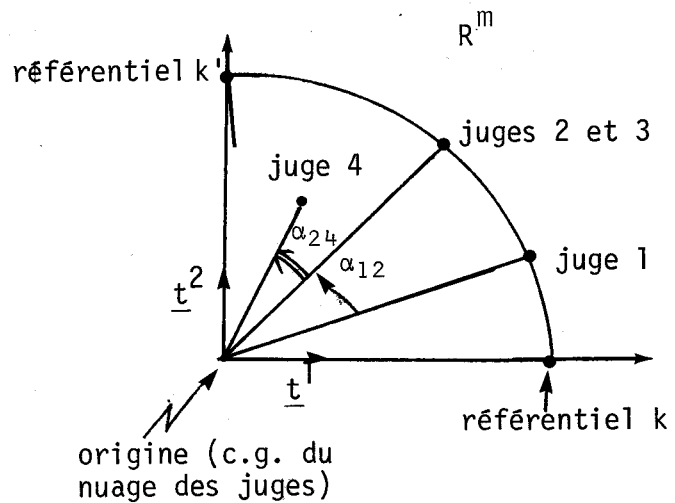
$$E_{kk'} = \text{Tr} (S_k S_{k'}) = RV (X_k, X_{k'}).$$

Deux juges qui appartiennent au plan (k, k') de l'inter-structure se trouvent sur le cercle de centre l'origine (centre de gravité du nuage des juges) et de rayon unité.

Les référentiels k et k' , normalisés, sont sur ce cercle.



avant normalisation



après normalisation

$$R^m \subset R^{n(n+1)/2}$$

Espace des juges

La proximité de deux juges apparaissant sur le cercle, 1 et 2 par exemple, est exactement représentée:

$$\cos \beta_{12} = RV(X_1, X_2).$$

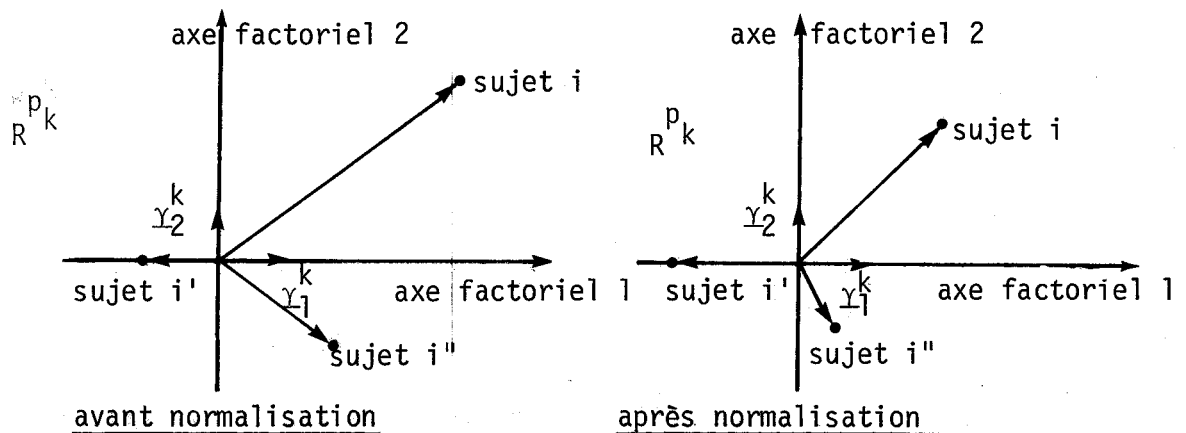
La proximité de deux juges ne se trouvant pas sur le cercle, 2 et 4 par exemple, n'est pas correctement représentée:

$$\cos \beta_{24} \neq RV(X_2, X_4).$$

Effet de la normalisation sur la k^e intra-structure

Les factorisations de S_k et $S_k / \|S_k\|$ fournissent deux représentations homothétiques (homothétie de centre l'origine et de facteur $\sqrt{\|S_k\|}$).

Par exemple, dans le premier plan factoriel de l'intra-structure:



Ceci montre que deux juges colinéaires dans l'inter-structure fournissent deux intra-structures homothétiques.

VII L'INTRA-STRUCTURE DE REFERENCE (COMPROMIS)

Les m matrices S_k du cube standard $S = \{S_k: n \times n; k = 1, \dots, m\}$, factorisées canoniquement, fournissent m représentations canoniques des sujets: $\{\gamma^k: n \times p_k; k = 1, \dots, m\}$ dans R^{p_k} (voir ci-dessus, Ch. IV, 5).

Chaque représentation est décrite dans un système particulier d'axes orthogonaux $\{\gamma_j^k; j = 1, \dots, p_k \leq n\}$ contenus dans R^{p_k} , les systèmes d'axes diffèrent d'une représentation à l'autre.

On cherche à comparer deux intra-structures, i.e., deux configurations $\boxed{N_k}$ et $\boxed{N_{k'}}$ (nuages des n sujets fournis par les juges k et k') autrement qu'en superposant les axes respectifs de chaque repère (la superposition directe ne constituant pas une bonne comparaison; [cf. L'Hermier (1976), p. 39].

On possède une intra-structure pour chaque juge et on voudrait une seule intra-structure qui soit une espèce de compromis entre les juges. L'intra-structure correspondant au k^e juge est obtenue en effectuant une factorisation canonique (ou ACP) sur la matrice S_k . On définit alors une matrice compormis comme étant une combinaison linéaire des matrices S_k . L'intra-structure compromis est obtenue en effectuant une factorisation canonique (ou ACP) sur cette nouvelle matrice.

Remarque: Le juge de référence

Disposant de deux populations,

- l'une de sujets sur laquelle est définie une proximité;
- l'autre de juges qui définit cette proximité;

on fait l'hypothèse qu'il existe un "juge de référence" par rapport à qui l'analyse des différences, dans les proximités données, est intéressante pour la compréhension de chaque jugement (i.e., la matrice de proximités d'un juge).

En général, ce juge de référence n'appartient pas à la population des juges donnés, mais son jugement est construit à partir des jugements de chacun des juges. (Le cas d'une matrice de référence à priori est exclu ici). Cette non appartenance est traduite par les termes de "matrice (de jugement) hypothétique" ou simplement par "matrice de référence".

1. Définitions de la matrice de référence (compromis)

Définition 1

Etant donné le cube standard $S = \{S_k: n \times n; k = 1, \dots, m\}$ et le premier vecteur propre $\underline{t}^1$ de $E: m \times m$, d'élément (k, k') égal à $\text{Tr}(S_k S_{k'})$, la matrice de référence (compromis) proposée par L'Hermier (1976) est:

$$r^1_{n \times n} = \sum_{k=1}^m t_k^1 S_k = \left(\sum_{k=1}^m t_k^1 s_{ii'}^k \right) = \left(t_1^1 s_{ii'}^1 + t_2^1 s_{ii'}^2 + \dots + t_m^1 s_{ii'}^m \right)$$

$$r^1 = \begin{bmatrix} \sum_{k=1}^m t_k^1 s_{11}^k & \cdots & \sum_{k=1}^m t_k^1 s_{1n}^k \\ \vdots & & \vdots \\ \sum_{k=1}^m t_k^1 s_{n1}^k & \cdots & \sum_{k=1}^m t_k^1 s_{nn}^k \end{bmatrix}$$

où t_k^1 désigne la k^e composante de $\underline{t}^1$ (voir ci-dessus, Ch. IV, 3).

Définition 2

Etant donné le premier vecteur référentiel $\underline{r}^1$, défini antérieurement (voir Ch. IV, 5), la matrice de référence r^1 est la matrice qui donne $\underline{r}^1$ par l'application $j_{i\ell} = \frac{i(i-1)}{2} + \ell$; $i = 1, \dots, n$; $\ell \leq i$.

2. Propriétés de la matrice de référence

- Les définitions 1 et 2 de la matrice de référence r^1 sont équivalentes.

Pour les démonstrations, voir L'Hermier (1976) [p. 43].

- Propriétés d'optimalité

La matrice r^1 maximise $\sum_{k'=1}^m \left[\langle s_{k'}, r^1 \rangle_{Tr} \right]^2$ parmi toute matrice r^k de la forme :

$$r^k = \sum_{k'=1}^m t_{k'}^k s_{k'}^k = \left(\sum_{k'=1}^m t_{k'}^k s_{ii'}^k \right)$$

sous la contrainte $\underline{t}^k \underline{t}^k = \text{constante} > 0$, le maximum étant (τ_2^1) pour $\underline{t}^1 \underline{t}^1 = < \underline{t}^1, \underline{t}^1 >_{R^{n(n+1)/2}} = 1$. Ainsi, r^1 est meilleure combinaison linéaire (dans un certain sens) des S_k .

Pour la démonstration, voir ci-dessus, Ch. IV, 4 et L'Hermier (1976) [pp 15-17].

. La matrice r^1 est définie positive

Pour la démonstration L'Hermier (1976) [pp. 44-45].

Remarque: Ce résultat assure la qualité de l'analyse des différences faite en se servant du premier référentiel comme juge hypothétique puisque la matrice r^1 se factorise exactement.

3. Obtention de l'intra-structure de référence

Considérons la factorisation canonique de r^1 ($n \times n$):

$$r^1 = a a^T$$

$$r^1 = \sum_{\ell=1}^n \underline{a}_{\ell} \underline{a}_{\ell}^T = \sum_{\ell=1}^n \sum_{\ell'=1}^n \alpha_{\ell} \underline{a}_{\ell\ell'} \underline{a}_{\ell\ell'}^T$$

où $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n > 0$ sont les valeurs propres de r^1 et $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$ sont les vecteurs propres normalisés: $\underline{a}_{\ell}^T \underline{a}_{\ell} = 1, \forall \ell = 1, \dots, n$ et $A = \begin{bmatrix} \alpha_1 & & 0 \\ & \alpha_2 & \\ 0 & & \ddots \\ & & & \alpha_n \end{bmatrix}$

La factorisation canonique de r^1 fournit la représentation canonique de référence à partir de laquelle est obtenue l'intra-structure de référence

(ou intra-structure compromis). Cette dernière est constituée de l'ensemble des projections sur les plans factoriels, définis par deux axes factoriels $\underline{a}_\ell$ et $\underline{a}_{\ell'}$, contenus dans R^n (espace compromis) et ordonnés selon le pourcentage

$$\frac{\alpha_\ell + \alpha_{\ell'}}{\sum_{\ell=1}^n \alpha_\ell}$$

de la variance expliquée. La meilleure représentation plane est celle dans $(\underline{a}_1, \underline{a}_2)$, au sens de la variance maximum expliquée. D'autres plans peuvent être considérés au besoin.

VIII. L'ANALYSE DES DIFFERENCES

1. Calcul des différences entre sujets

La k^e représentation canonique des sujets dans R^{p_k} est définie (voir ci-dessus, Ch. IV, 5) par:

$$\gamma^k: n \times p_k; k = 1, \dots, m; p_k \leq n$$

où $\{\lambda_k^j, \gamma_j^k; j = 1, \dots, p_k \leq n\}$ est l'ensemble des éléments propres de S_k . La k^e intra-structure représente dans R^{p_k} (ou dans quelques plans) les n sujets tels que vus par le k^e juge.

La représentation canonique de la matrice de référence (ou compromis) est définie (voir ci-dessus, Ch. VII, 2) par:

$$a_{n \times n} = (\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n)$$

où $\{\alpha_\ell, \underline{a}_\ell; \ell = 1, \dots, q = n\}$ est l'ensemble des éléments propres de r^1 . Les vecteurs propres $\underline{a}_\ell$ constituent une base dans l'espace des variables $R^q = R^n$. L'intra-structure de référence (ou intra-structure compromis) représente aussi dans R^n (ou dans quelques plans) les n sujets tels que vus par un juge compromis.

Les bases des deux espaces précédents sont évidemment différentes puisqu'elles sont définies par des vecteurs propres normalisés correspondant à des matrices différentes.

La base de la k^e intra-structure est donnée par $\underline{y}_1^k, \underline{y}_2^k, \dots, \underline{y}_{p_k}^k$, qui sont les vecteurs propres de S_k normalisés à 1 en ordre décroissant des valeurs propres. La base de l'intra-structure de référence est $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$, qui sont les vecteurs propres normalisés de r^1 à 1, en ordre décroissant des valeurs propres.

Si $\underline{x}_i^k$ ($n \times 1$) représente le i^e sujet dans la k^e intra-structure, alors on a:

$$\underline{x}_i^k = \sum_{j=1}^{p_k \leq n} x_{ij}^k \underline{y}_j^k, \quad i = 1, \dots, n.$$

Nous voulons représenter ce i^e sujet (tel que vu par le k^e juge) dans l'intra-structure de référence. Si $\underline{z}_i^k$ ($n \times 1$) le représente, alors on a

$$\underline{z}_i^k = \sum_{\ell=1}^n z_{i\ell}^k \underline{a}_\ell, \quad i = 1, \dots, n.$$

Le choix de $\underline{z}_i^k$ est tel que:

$$\| \underline{z}_i^k - \underline{x}_i^k \|^2 = \left(\underline{z}_i^k - \underline{x}_i^k \right)^T \left(\underline{z}_i^k - \underline{x}_i^k \right) \text{ soit minimum}$$

Cela sera fait pour tout sujet $i = 1, 2, \dots, n$ tel que vu par chacun des juges $k = 1, 2, \dots, m$.

On montre [Cf. L'Hermier (1976), p. 40-42] que $\| \underline{z}_i^k - \underline{x}_i^k \|^2$ est minimum lorsque:

$$z_{i\ell}^k = \frac{\sum_{j=1}^{p_k} x_{ij}^k < \underline{y}_j^k, \underline{a}_\ell >}{< \underline{a}_\ell, \underline{a}_\ell >}, \quad \ell = 1, \dots, n$$

Dans l'espace de référence (ou espace compromis) R^n , le produit scalaire $< \underline{y}_j^k, \underline{a}_\ell > = \text{corrélation}(\underline{x}_j, \underline{z}_\ell)$ tandis que $< \underline{a}_\ell, \underline{a}_\ell > = 1$; d'où l'expression simplifiée de $z_{i\ell}^k$:

$$z_{i\ell}^k = \sum_{j=1}^{p_k} x_{ij}^k \underline{y}_j^k \underline{a}_\ell \quad \ell = 1, \dots, n$$

La matrice, ayant pour $(j, \ell)^e$ élément le terme $\langle \gamma_j^k, a_\ell \rangle$ est identique à la matrice ayant pour $(j, \ell)^e$ élément le terme

$$\frac{\text{Cov}(\underline{x}_j, \underline{z}_\ell)}{\sqrt{\text{Var}(\underline{x}_j) \text{Var}(\underline{z}_\ell)}} = \text{Cor}(\underline{x}_j, \underline{z}_\ell)$$

Par conséquent, le vecteur représentatif de tout sujet i ($i = 1, \dots, n$) dans l'espace compromis R^n (des sujets) est donné par

$$\begin{matrix} \underline{z}_i \\ (1 \times n) \end{matrix} = \begin{matrix} \underline{x}_i \\ (1 \times p_k) \end{matrix} \begin{matrix} \sigma_{11}^{-\frac{1}{2}} & \sigma_{12} & \sigma_{22}^{-\frac{1}{2}} \end{matrix}$$

en posant

$$\begin{matrix} \sigma_{11} \\ (p_k \times p_k) \end{matrix} = \text{diag} \left[\text{Var}(\underline{x}_j); j = 1, \dots, p_k \right] = \begin{bmatrix} \text{var}(\underline{x}_1) & & 0 \\ & \text{var}(\underline{x}_2) & \\ 0 & & \text{var}(\underline{x}_{p_k}) \end{bmatrix}$$

$$\begin{matrix} \sigma_{22} \\ (n \times n) \end{matrix} = \text{diag} \left[\text{Var}(\underline{z}_\ell); \ell = 1, \dots, n \right] = \begin{bmatrix} \text{var}(\underline{z}_1) & & 0 \\ & \text{var}(\underline{z}_2) & \\ 0 & & \text{var}(\underline{z}_n) \end{bmatrix}$$

$$\begin{matrix} \sigma_{12} \\ (p_k \times n) \end{matrix} = \text{Cov}(\underline{x}_j, \underline{z}_\ell) = \begin{bmatrix} \text{Cov}(\underline{x}_1, \underline{z}_1) & \dots & \text{Cov}(\underline{x}_1, \underline{z}_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(\underline{x}_{p_k}, \underline{z}_1) & \dots & \text{Cov}(\underline{x}_{p_k}, \underline{z}_n) \end{bmatrix}$$

La configuration Z : $p_k \times n$ des n sujets dans R^n la plus proche de la configuration X : $p_k \times n$ est donnée par:

$$Z = X \sigma_{11}^{-\frac{1}{2}} \cdot \sigma_{12} \cdot \sigma_{22}^{-\frac{1}{2}}$$

ou bien par:

$$Z' = \sigma_{22}^{-\frac{1}{2}} \cdot \sigma_{21} \cdot \sigma_{11}^{-\frac{1}{2}} \cdot X' \quad \text{avec} \quad \sigma_{21} = \sigma_{12}'$$

(n x p_k)

Par contre, dans l'espace R^{p_k} des sujets, chacune des représentations canoniques γ^k : n x p_k des sujets est exprimée dans le repère (base) de la représentation de référence, encore appelé repère référentiel.

Dans ce cas là, on a

$$\sigma_{22} = (\underline{a}_\ell, \underline{a}_\ell) = a^T a = \text{diag} (\alpha_1, \alpha_2, \dots, \alpha_n)$$

(n x n)

ou encore

$$\sigma_{22}^{-1} = (a^T a)^{-1} = \text{diag} (1/\alpha_1, 1/\alpha_2, \dots, 1/\alpha_n) =$$

(n x n)

$$\begin{bmatrix} 1/\alpha_1 & & 0 \\ & 1/\alpha_2 & \\ 0 & & \ddots & \\ & & & 1/\alpha_n \end{bmatrix}$$

et

$$\sigma_{21}^k = \left(\underline{a}_\ell, \underline{\gamma}_j^k \right) = a^T \gamma^k = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} \gamma_{11}^k & \dots & \gamma_{1p_k}^k \\ \vdots & & \vdots \\ \gamma_{n1}^k & \dots & \gamma_{np_k}^k \end{bmatrix}$$

(n x p_k) (n x n) (n x p_k)

d'où l'on en déduit la représentation canonique $\hat{a}^k$:

$$\hat{a}^k = \begin{matrix} (n \times n) & \left[\begin{array}{ccc} \hat{a}_{11}^k & \text{---} & \hat{a}_{1i}^k & \text{---} & \hat{a}_{1n}^k \\ \vdots & & \vdots & & \vdots \\ \hat{a}_{i1}^k & \text{---} & \hat{a}_{ii}^k & \text{---} & \hat{a}_{in}^k \\ \vdots & & \vdots & & \vdots \\ \hat{a}_{n1}^k & \text{---} & \hat{a}_{ni}^k & \text{---} & \hat{a}_{nn}^k \end{array} \right] \end{matrix}, \forall k = 1, \dots, m.$$

Donc, tous les juges peuvent représenter tous les sujets dans l'espace du juge compromis. Si tous les sujets sont bien représentés dans leur premier plan factoriel respectif, alors on peut visualiser dans le plan $(\underline{a}_1, \underline{a}_2)$ par exemple, les différences de représentation entre les juges et entre le juge compromis.

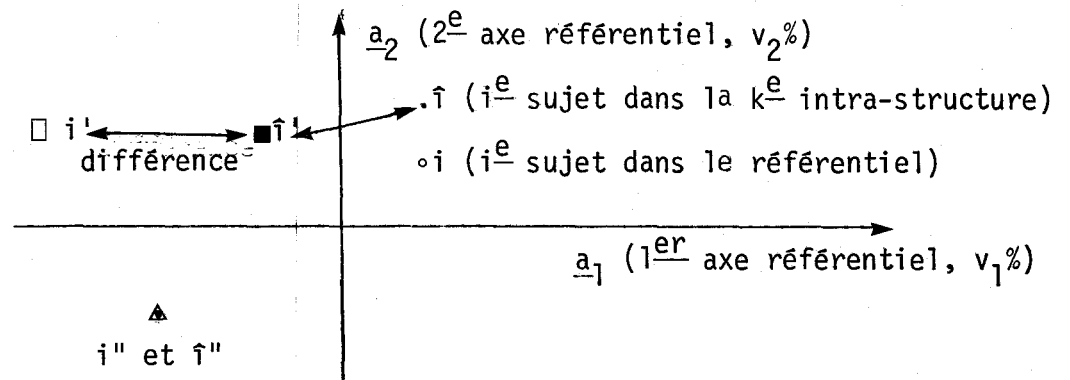
2. Visualisation des différences

De la représentation canonique $\hat{a}^k$ (matrice ayant pour i^e ligne $(\hat{a}_{i1}^k, \hat{a}_{i2}^k, \dots, \hat{a}_{ii}^k, \dots, \hat{a}_{in}^k)$), on extrait les plans $(\underline{\hat{a}}_1^k, \underline{\hat{a}}_2^k)$, $(\underline{\hat{a}}_1^k, \underline{\hat{a}}_3^k)$, ... etc de la k^e intra-structure ($k = 1, \dots, m$) rapportée au repère référentiel $(\underline{a}_\ell, \underline{a}_{\ell'})$.

On visualise simultanément autant de couples de plans qu'il est nécessaire pour expliquer un pourcentage de variance fixé à priori:

$$\left((\underline{\hat{a}}_1^k, \underline{\hat{a}}_2^k) \text{ et } (\underline{a}_1, \underline{a}_2) \right) \text{ puis } \left((\underline{\hat{a}}_1^k, \underline{\hat{a}}_3^k) \text{ et } (\underline{a}_1, \underline{a}_3) \right) \dots \text{ etc.}$$

Exemple: Visualisation des différences entre sujets dans le premier couple de plans:



Ainsi, les différences de positionnement sont visualisées; les pourcentages de variance expliquée par chaque axe, soit v_ℓ pour le nuage de référence et v_ℓ^k pour le nuage du k^e juge, fournissent une indication sur la pondération globale et relative de l'axe ℓ par le k^e juge:

$$\frac{v_\ell^k}{v_\ell} = \text{pois donné par le } k^e \text{ juge à l'axe } \ell$$

(indication approximative)

3. Mesure de la qualité des représentations

Le coefficient RV

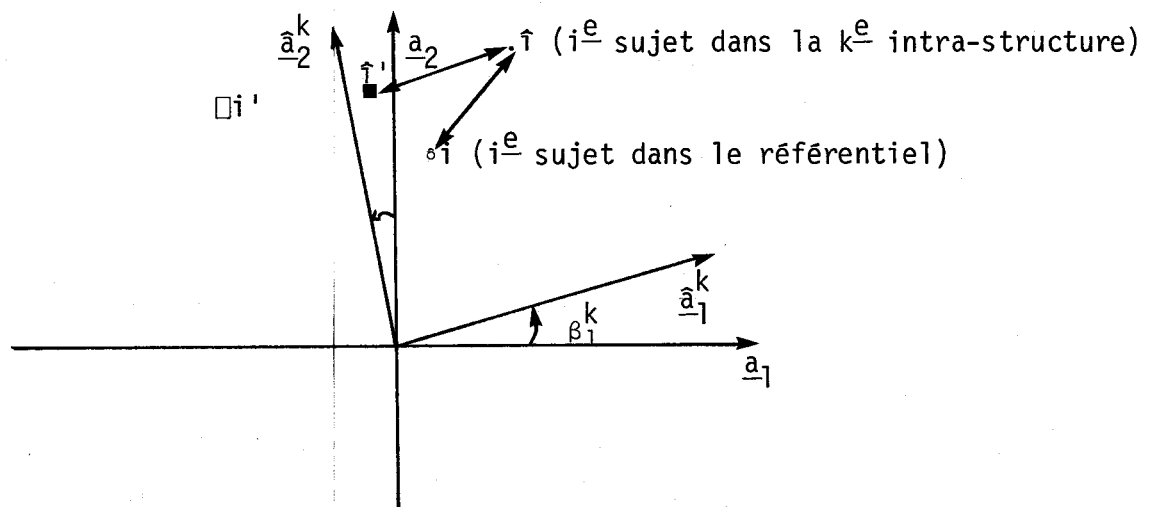
Dans la k^e intra-structure, les n sujets forment un nuage de points dans R^{pk} . Dans l'intra-structure compromis, ce nuage est aussi représenté dans R^n . Les projections des nuages dans le repère référentiel introduisent des dis-

torsions quant aux distances visualisées entre sujets. Le calcul du coefficient RV entre ces deux nuages donne une mesure de la qualité de l'approximation dans l'intra-structure compromis. Le coefficient RV entre l'intra-structure de référence et la k^e intra-structure est égal à:

$$RV(a, \gamma^k) = \frac{\text{Tr}(a a^T \gamma^k (\gamma^k)^T)}{\|a a^T\| \|\gamma^k (\gamma^k)^T\|} = \frac{\text{Tr}((\gamma^k)^T a a^T \gamma^k)}{\|r^1\| \|s^k\|} = \frac{\text{Tr}(\sigma_{21}^T \sigma_{12})}{\|r^1\| \|s^k\|}$$

Le coefficient RV n'indique pas directement le pourcentage de variance expliquée après projection; en effet, on peut très bien avoir simultanément, pour la k^e intra-structure, un pourcentage v^k égal avant et après projection et $RV(a, \gamma^k) \neq 1$. Cela tient au fait que le coefficient RV mesure (par rapport à l'intra-structure de référence a) non seulement une variation dans la disposition des sujets décrits dans la k^e intra-structure, mais aussi une variation dans l'orientation des axes de la k^e intra-structure.

Exemple: Visualisation des variations dans la disposition des sujets et dans l'orientation des axes



Dans la mesure où σ_{21} détermine la projection du nuage de points relatifs aux n sujets dans le système d'axes déduits de a et γ^k , le coefficient RV entre l'intra-structure de référence a et la k^e intra-structure γ^k donne une indication globale de la distorsion due à la projection de γ^k sur a .

Cas particuliers

RV (a, γ^k) = 1 lorsque:

- $Z_k = X_k$, c'est-à-dire $a = \gamma^k$; d'où

$$\begin{aligned} \sigma_{21}^k &= a^T \gamma^k = a^T a = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \text{ et comme} \\ (\sigma_{22}^k)^{-1} &= (a^T a)^{-1} = \text{diag}(1/\alpha_1, 1/\alpha_2, \dots, 1/\alpha_n), \text{ alors} \\ (\sigma_{22}^k)^{-1} \sigma_{21}^k &= I_n \quad \text{et} \quad \hat{a}^k = \gamma^k. \end{aligned}$$

- $Z_k = X_k$ à une permutation circulaire près, i.e. une rotation orthogonale, près des axes.

Pour la démonstration, voir L'Hermier (1976) [p. 48].

RV (a, γ^k) = 0 lorsque:

- $a = 0$, ce qui implique $S_k = 0$, $\forall k = 1, \dots, m$, car alors

$$E = 0 \implies \underline{t}^1 = \underline{0} \quad \text{et} \quad r^1 = 0.$$

$$\gamma^k = 0$$

Dans ces deux derniers cas, l'analyse des différences est sans intérêt.

Calcul du pourcentage de variance visualisée

Une indication plus précise de la distorsion est fournie en recalculant le pourcentage $v_{\ell\ell}^k$ de variance expliquée par chaque plan $(\hat{a}_{\ell}^k; \hat{a}_{\ell}^k)$ de la k^e intra-structure après projection ($v_{\ell\ell}^k$ correspondant au pourcentage de la variance expliquée avant projection).

IX. L'ANALYSE DES EVOLUTIONS

Lorsque le 3^e indice, k , du cube standard $S = \{S_k: n \times n; k = 1, \dots, m\}$ marque des époques dans le temps, c'est-à-dire, lorsque S_k représente le "point de vue" d'un juge à un instant donné et S , le "point de vue" de ce même juge à des instants différents, S est dit chronologique.

L'analyse des différences devient alors une analyse de l'évolution des intra-structures.

1. Détermination des intra-structures dans le référentiel (compromis)

Chaque représentation canonique $\gamma^k: n \times p_k$ ($k = 1, \dots, m$) des sujets dans chaque intra-structure est donc exprimée dans le repère référentiel ou intra-structure compromis, comme il est fait dans l'analyse des différences.

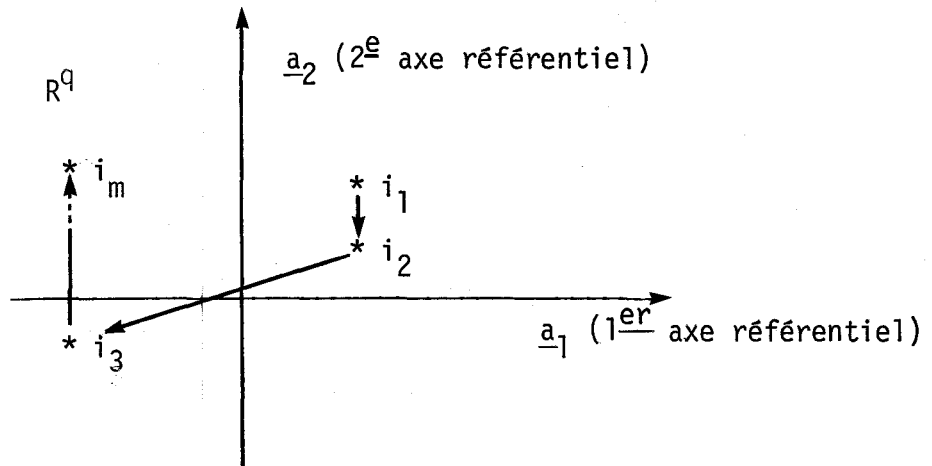
On obtient m représentations canoniques $\hat{a}^k: n \times n$ ($k = 1, \dots, m$) des sujets dans R^n rapportées au même repère référentiel.

2. Visualisation de l'évolution

Les plans (ℓ, ℓ') , pris dans l'ordre des pourcentages décroissants de la variance expliquée, de toutes les intra-structures $\{(\hat{a}_{\ell}^k, \hat{a}_{\ell'}^k); k = 1, \dots, m\}$ sont simultanément visualisés.

Le pourcentage de variance expliquée par chaque axe ℓ ou par chaque plan (ℓ, ℓ') , soit v_{ℓ}^k ou $v_{\ell\ell'}^k$, ($k = 1, \dots, m$), est alors donné pour les m configurations.

Exemple: Positionnement, noté i_k , du sujet i ($i = 1, \dots, n$) dans le plan $(\hat{a}_1^k, \hat{a}_2^k)$ de la k^{e} configuration ($k = 1, \dots, m$) et rapporté au premier plan référentiel $(\underline{a}_1, \underline{a}_2)$.



Ainsi, dans l'intra-structure de référence (ou compromis), chaque sujet est représenté m fois, une pour chaque instant. Lorsque la représentation peut se faire adéquatement dans le premier plan factoriel de référence $(\underline{a}_1, \underline{a}_2)$, on obtient alors une représentation visuelle de l'évolution de chaque sujet.

3. Mesure de la qualité des représentations

Comme il est fait dans l'analyse des différences, le coefficient RV (a, γ^k) donne, pour chaque représentation canonique γ^k ($k = 1, \dots, m$) une indication globale de la distorsion due à la projection de γ^k sur a .

La répartition de cette distorsion selon les plans $(\underline{a}_\ell^k, \underline{a}_{\ell'}^k)$ est donnée par les pourcentages de la variance expliquée et visualisée par les axes $\underline{a}_\ell^k$ ($\ell = 1, \dots, n; k = 1, \dots, m$) définissant ces plans.

X. EXTENSIONS DE LA METHODE STATIS

1. Autres matrices de référence

Définition

La définition, donnée pour la première matrice de référence r^1 , peut être généralisée.

Etant donnés les éléments propres $\{\tau_k, \underline{t}^k; k = 1, \dots, m\}$ ou spectre de la matrice $E: m \times m$ d'élément $E_{kk'} = \text{Tr}(S_k S_{k'})$ où $S_k \in S = \{S_k: n \times n; k = 1, \dots, m\}$.

Une k'^e matrice de référence ($k' = 1, \dots, m$) peut être définie par:

$$r^{k'} = \sum_{k=1}^m t_k^{k'} S_k$$

où $t_k^{k'}$ désigne la k^e composante du vecteur propre $\underline{t}^{k'}$ (voir ci-dessus, Ch. IV, 3).

Propriétés

Toutes ces matrices $r^{k'}$ ($k' = 1, \dots, m$) vérifient:

- . la propriété d'optimalité (voir ci-dessus, Ch. VII, les maximums étant $(\tau_{k'})^2$ respectivement.
- . $r^{k'}$ est "matrice propre" de VoM qui applique $\mathcal{F}(R^n) \rightarrow \mathcal{F}(R^n)$ dans le schéma naturel [Cf. L'Hermier (1976), p. 13].
- . par contre, la propriété de définie-positivité (voir ci-dessus, Ch. VII, 1) n'est assurée que pour $k' = 1$, ce qui enlève quelque peu d'intérêt à ces matrices de référence $r^2, \dots, r^m$ pour les analyses des différences ou de l'évolution. Car pour ces analyses, les matrices $r^{k'}$,

$k' = 2, \dots, m$, doivent être rendues définies-positives; ce qui a pour conséquence, outre l'approximation, que la propriété d'optimalité n'est plus vérifiée par ces approximations $\hat{r}^{k'}$; on a seulement:

$$0 \leq \sum_{k'=1}^m \left[\langle s_k, \hat{r}^{k'} \rangle_M \right]^2 \leq \tau_k^2.$$

De plus, $\hat{r}^{k'}$ ne sont plus "matrices propres" et, par conséquent, il n'est plus possible de les positionner dans l'inter-structure (voir, ci-dessus, Ch. IV, 5). Il faudrait faire appel à la technique décrite par Gower (1968) pour réaliser correctement ce positionnement.

2. Autres structures

Les structures (inter ou intra) proposées découlent des choix essentiels suivants:

- des métriques M et N dans le schéma de dualité proposé;
- de l'usage de la factorisation canonique (ou ACP).

3. Tableaux à q indices, $q > 3$

Lorsque $q = 4$, la donnée est un ensemble de cubes tels que:

$$H = \{S^1, \dots, S^r\} = \{S^\ell: n \times n \times m; \ell = 1, \dots, r\};$$

un juge est un cube $S^\ell: n \times n \times m$ ($\ell = 1, \dots, r$) tandis qu'un sujet est une matrice de produits scalaires entre matrices de proximités (ou de profils) $S_k: n \times n$ ($k = 1, \dots, m$)

L'inter-structure visualise les cubes $S^1, \dots, S^r$ tandis que les intra-structures sont les inter-structures issues de chacun des cubes.

REFERENCES

CAILLIEZ, F. et J.P. PAGES. (1976).

Introduction à l'analyse des données. SMASH, Paris 1976, 616 pages.

CLEROUX, R.

Analyse de données multidimensionnelles. Notes de cours, D.I.R.O.

Université de Montréal, p. 248-269.

ESCOUFIER, Y. (1970).

Echantillonnage dans une population de variables aléatoires réelles.

Pub. Inst. Stat. Univ. Paris, Vol. 19 (4): 1-47.

ESCOUFIER, Y. (1973).

Le traitement des variables vectorielles. Biometrics, 29 (4): 750-760.

ESCOUFIER, Y. (1975).

Le positionnement multidimensionnel. Revue de Statistique Appliquée

23 (4): 5-14.

ESCOUFIER, Y. (1977a).

Où en est-on dans l'utilisation des opérateurs? Journée d'étude sur

l'Analyse des données, Paris, avril 1977.

ESCOUFIER, Y. (1977b).

Operators related to a data matrix. In: Recent developments in Statistics, J.R. Barra et al., editors; North-Holland, 1977; p. 125-131.

ESCOUFIER, Y. , F. CAILLIEZ et J.P. PAGES. (1978).

Géométrie et techniques particulières en analyse factorielle.

Rapport technique, Université de Montpellier, mai 1978.

ESCOUFIER, Y. et H. L'HERMIER. (1978).

A propos de la comparaison graphique des matrices de variance.

Biom. J. 20 (5): 491-497.

ESCOUFIER, Y. et P. ROBERT. (1977).

Choosing variables and metrics by optimizing the RV-coefficient.

In: RUSTAGI, J.S. (ed.), Proceedings of the Second International Conference on Optimisation in Statistics, Bombay, dec. 1977; Academic Press.

GOWER, J.C. (1968).

Some distances properties of latent root and vector methods used in multivariate analysis. Biometrika, 53 (3 et 4): 325-338.

L'HERMIER, H. (1976).

Structuration des tableaux à trois indices de la statistique. Thèse de 3^e cycle, Univ. Sci. Tech. Languedoc, Montpellier; 98 pages.

L'HERMIER, H. et B. THIEBAUT. (1977).

Etude de la pluviométrie au moyen de la méthode STATIS. Revue de Statistique Appliquée, 25: 57-81.

PAGES, J.P., Y. ESCOUFIER et P. CAZES. (1976).

Opérateurs et analyse des tableaux à plus de deux dimensions. Cahiers du BUR0, série Recherche, No 25, p. 62-89.

ROBERT, P. et Y. ESCOUFIER. (1976).

A Unifying Tool for linear multivariate statistical methods: the RV-coefficient. J. R. Stat. Soc., séries C (Appl. Stat.), 25 (3): 257-265.

TORGERSON, W.S. (1958).

Theory and method of scaling. Wiley, New-York, 1958.

ANNEXE A

	Page
A1 Schéma de dualité pour les analyses de données cubiques	A1
A2 Représentation du juge S_k dans le schéma de dualité	A4
A3 Schéma de dualité associé à l'ACP de la matrice X_k	A5
A4 Approximation d'une matrice, au sens des moindres carrés généralisés	A6
A5 Schéma de dualité associé au k^e vecteur référentiel	A6
A6 Calcul de la matrice de référence r^1	A8

A1. Schéma de dualité pour les analyses de données cubiques

Proposition

Soit $S = \{S_k: n \times n; k = 1, \dots, m\}$ le cube standard déduit des données.

$S_k \in \mathcal{P}(R^n)$, espace vectoriel des formes quadratiques sur R^n .

Soit x_k , le vecteur-colonne de $R^{n(n+1)/2}$ construit à partir de S_k par l'application, notée $j_{i\ell}$, et définie par:

$$x_{j_{i\ell}}^k = s_{i\ell}^k = \begin{bmatrix} s_{11} & 0 \\ s_{21} & s_{22} \\ s_{31} & s_{32} & s_{33} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \text{ et } j_{i\ell} = \frac{i(i-1)}{2} + \ell; i = 1, \dots, n; \ell \leq i,$$

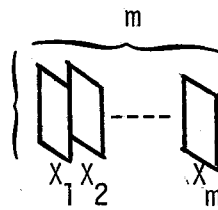
avec

$$S_k = (s_{i\ell}^k) = \begin{bmatrix} s_{11}^k & & 0 \\ \vdots & s_{22}^k & \\ s_{n1}^k & \dots & s_{nn}^k \end{bmatrix} \text{ et } x_k = (x_{j_{i\ell}}^k) = \begin{bmatrix} x_{11}^k \\ x_{21}^k \\ x_{22}^k \\ x_{31}^k \\ x_{32}^k \\ x_{33}^k \\ \vdots \end{bmatrix}$$

et $x_{j_{i\ell}}^k$ est la "mesure-produit" de la $j_{i\ell}^e$ variable (caractère) faite sur les individus i et ℓ par le k^e juge.

Soit X : $n(n+1)/2 \times m$, la matrice qui admet X_k pour k^{e} colonne

$$X = [X_1 \ X_2 \ \dots \ X_k \ \dots \ X_m]$$

$\frac{n(n+1)}{2}$


Enfin, soit M : $n(n+1)/2 \times n(n+1)/2$, la matrice de transformation diagonale, définie par:

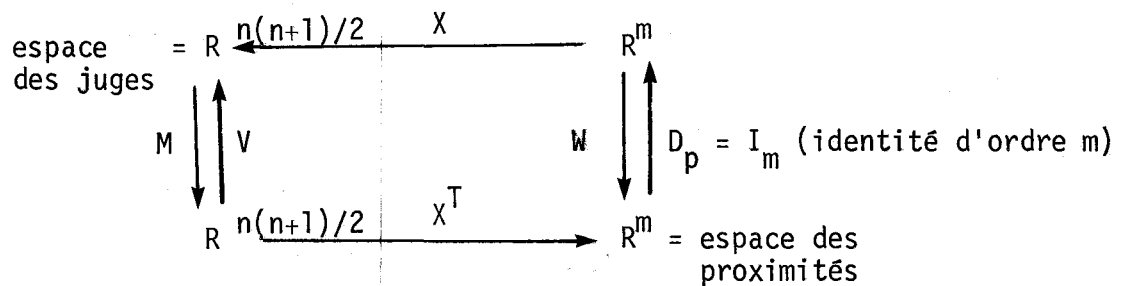
$$m_{j_{il}} = \begin{cases} 2 & \text{si } i < l \\ 1 & \text{si } i = l \end{cases}$$

c'est-à-dire

$$M = (m_{j_{il}}) = \begin{bmatrix} 1 & & & & 0 \\ & 2 & & & \\ & & 1 & & \\ & & & 2 & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \quad \left. \vphantom{\begin{bmatrix} 1 & & & & 0 \\ & 2 & & & \\ & & 1 & & \\ & & & 2 & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix}} \right\} n(n+1)/2$$

m

Le schéma de dualité proposé par L'Hermier (1976) est alors:



Propriété de ce schéma de dualité

La distance mesurée entre deux juges, X_k et $X_{k'}$, vecteurs de $R^{n(n+1)/2}$ est égale à la distance entre les deux matrices, S_k et $S_{k'}$, de $\mathcal{F}(R^n)$ correspondant à ces vecteurs, lorsqu'on prend la métrique Trace:

$$\langle X_k, X_{k'} \rangle_M = \langle S_k, S_{k'} \rangle_{\text{Trace}}; \quad X_k \xrightleftharpoons{j_{il}} S_k$$

Pour la démonstration, voir L'Hermier (1976) [p. 12].

Conséquence

La matrice associée à l'opérateur WD_p , c'est-à-dire, la matrice des proximités entre deux vecteurs de $R^{n(n+1)/2}$, X_k et $X_{k'}$, représentant respectivement le juge k et le juge k' , peut être calculée par la formule qui donne pour $(k, k')^E$ élément: $\text{Trace}(S_k S_{k'})$ puisque la donnée S_k est préalable à l'application j_{il} :

$$s_{il}^k \xrightleftharpoons{j_{il}} x_{j_{il}}^k$$

$$W = (\text{Tr}(S_k S_{k'})) = (\langle X_k, X_{k'} \rangle_M) = (\text{cov}(W_k D_p, W_{k'} D_p))$$

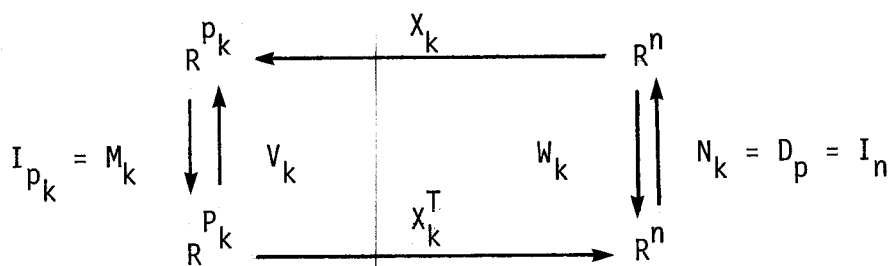
Remarque. [voir L'Hermier (1976), p. 13-14].

A la famille des tableaux de données $\{X_k, M_k, D_p; k = 1, \dots, m\}$ est associée la famille des opérateurs $\{W_k, D_p, k = 1, \dots, m\}$. Nous nous proposons d'étudier les proximités et les différences entre les études X_k .

A2. Représentation du juge S_k dans le schéma de dualité

La factorisation canonique de la matrice S_k fournit une matrice X_k : $(p_k \times n)$ telle que $S_k = X_k^T X_k$.

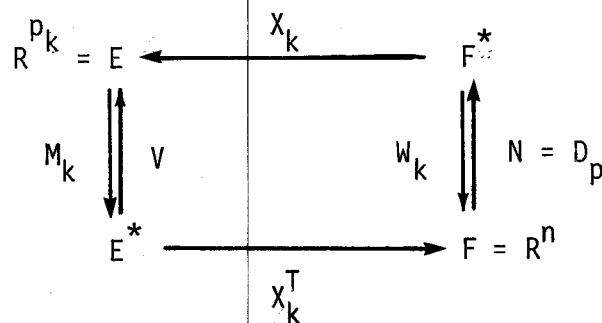
Si la matrice X_k est associée à l'application X_k dans le schéma de dualité (classique en ACP), alors la matrice S_k est associée à l'application W_k , la matrice identité d'ordre p_k , I_{p_k} , est associée à l'application M_k et la matrice identité d'ordre n , I_n , est associée à l'application $N_k = D_p$ dans le schéma de dualité:



Ceci montre comment le schéma proposé par L'Hermier (1976) est relié aux m schémas classiques (de l'ACP); cela permet alors de préciser la nature de la proximité entre deux juges.

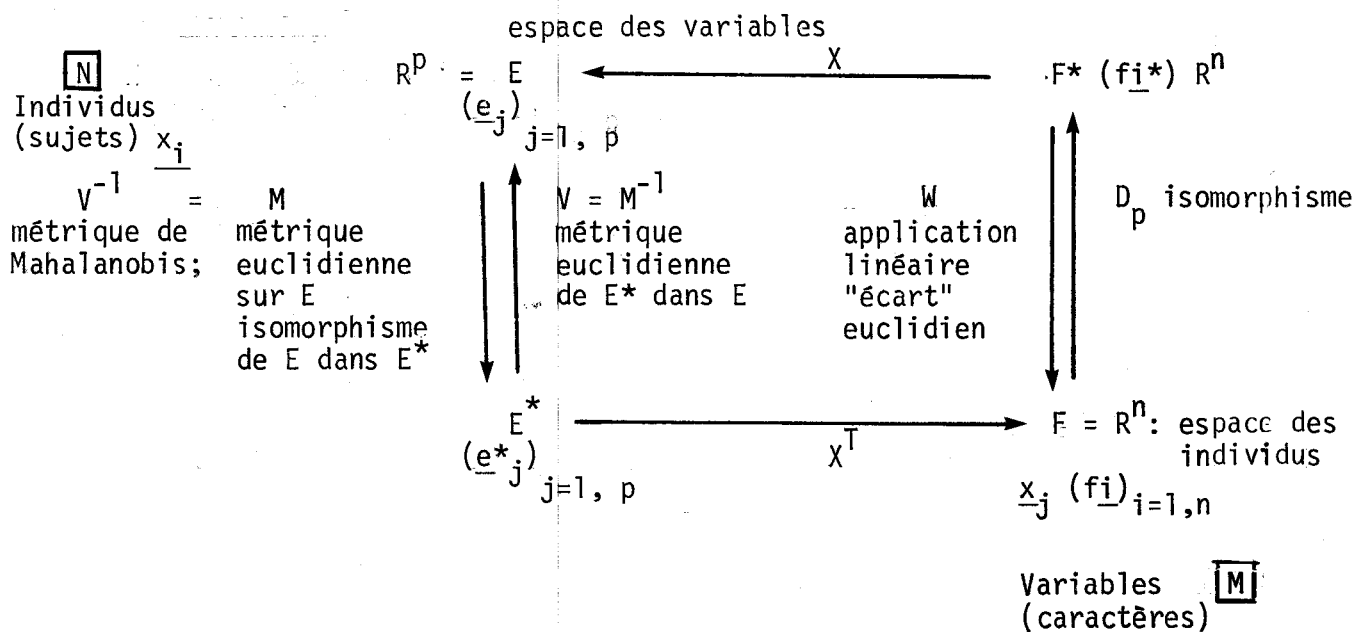
A3. Schéma de dualité associé à l'ACP de la matrice X_k

L'application $A_k = X_k^T M_k X_k$ n'est autre que l'application, notée W_k , du schéma de dualité associé à l'ACP de la matrice des données X_k :



Note: La transformation, notée T_1 dans Cléroux, (p. 254), n'est autre que cette application $A_k = W_k: F^* \longrightarrow F = R^n$.

Complément. Le schéma de dualité ci-dessus peut être repris avec plus de détail (sans spécifier l'indice "k", i.e. son appartenance au k^E juge pour simplifier la notation).



A4. Approximation d'une matrice, au sens des moindres carrés généralisés

Théorème

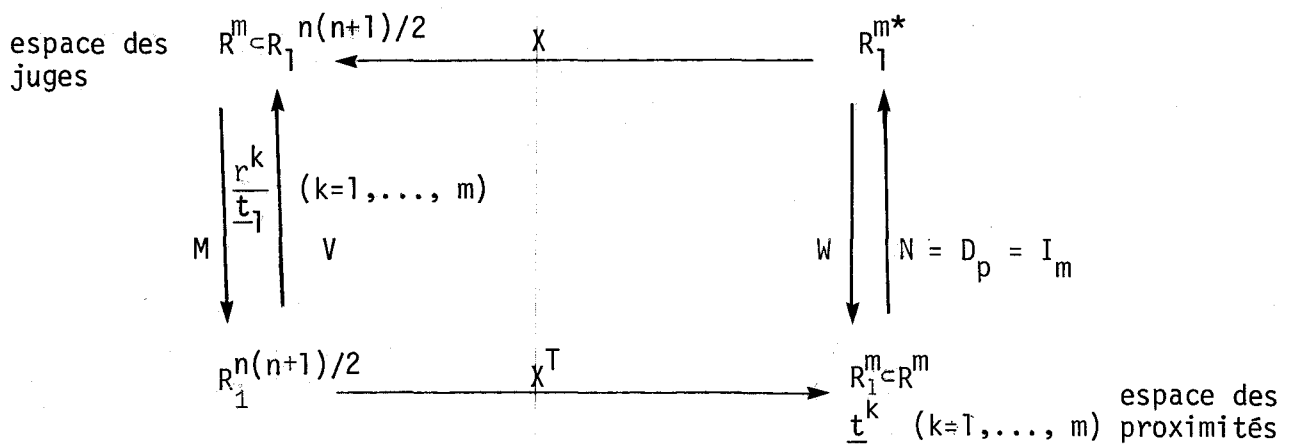
Si S est une matrice réelle symétrique et si $\text{Tr} (S - T^T T)^2$ est minimisée par rapport à une matrice réelle T , alors

$$T^T T = \sum_{\lambda_i > 0} \lambda_i u_i u_i^T$$

où λ_i , u_i sont respectivement les valeurs propres positives et leurs vecteurs propres associés de S .

Pour la démonstration, voir L'Hermier (1976) [p. 92-95: théorème 2].

A5. Schéma de dualité associé au k^e vecteur référentiel



$$\underset{m \times m}{t} = (\underset{1 \times m}{t^1}, \underset{1 \times m}{t^2}, \dots, \underset{1 \times m}{t^m}) = \begin{bmatrix} \underset{1 \times m}{t^1_1} & \underset{1 \times m}{t^2_1} & \dots & \underset{1 \times m}{t^k_1} & \dots & \underset{1 \times m}{t^m_1} \\ \vdots & \vdots & & \vdots & & \vdots \\ \underset{1 \times m}{t^1_k} & \underset{1 \times m}{t^2_k} & \dots & \underset{1 \times m}{t^k_k} & \dots & \underset{1 \times m}{t^m_k} \\ \vdots & \vdots & & \vdots & & \vdots \\ \underset{1 \times m}{t^1_m} & \underset{1 \times m}{t^2_m} & \dots & \underset{1 \times m}{t^k_m} & \dots & \underset{1 \times m}{t^m_m} \end{bmatrix} = \begin{bmatrix} \underset{1 \times m}{t_1} \\ \vdots \\ \underset{1 \times m}{t_2} \\ \vdots \\ \underset{1 \times m}{t_m} \end{bmatrix}$$

vecteur ligne $\underset{1 \times m}{t_k}$ ($1 \times m$): image du k^{e} juge dans R^m (espace des juges)

vecteur colonne $\underset{m \times 1}{t^k}$ ($m \times 1$): composantes des m juges sur le k^{e} axe contenu dans R^m .

$\underset{m \times 1}{t^1}$: composantes des m juges sur le 1^{er} axe factoriel contenu dans R^m (espace des proximités)

$$\underset{1 \times m}{r^k} = XN(\underset{m \times 1}{t^k}) = \sum_{k'=1}^m \underset{1 \times m}{t^k_{k'}} \underset{1 \times m}{x_{k'}} \quad \text{avec } (\underset{m \times 1}{t^k})^T \underset{m \times 1}{t^{k'}} = \delta_{kk'} \text{ (Kronecker)}$$

et $\underset{1 \times m}{x_{k'}}$: k'^{e} colonne de la matrice X ; $X = (\underset{1 \times m}{x_1}, \underset{1 \times m}{x_2}, \dots, \underset{1 \times m}{x_{k'}}, \dots, \underset{1 \times m}{x_m})$
 $n(n+1)/2 \times m$

$$\underset{1 \times m}{r^k} = \begin{bmatrix} \underset{1 \times m}{t^k_1} \\ \underset{1 \times m}{t^k_2} \\ \vdots \\ \underset{1 \times m}{t^k_{k'}} \\ \vdots \\ \underset{1 \times m}{t^k_m} \end{bmatrix} (x^{k'}_{jil})$$

A6. Calcul de la matrice de référence r^1 :

avec $E_{mxm} = \underline{t} \underline{t}^T$

$$E = \underline{t}^k T (\underline{t}^k)^T = \sum_{\substack{k'=1 \\ \tau_k > 0}}^m \tau_k \underline{t}_{k'}^k (\underline{t}_{k'}^k)^T$$

lorsque $(\underline{t}^k)^T \underline{t}^k = 1$ et

$$T = \begin{bmatrix} \tau_1 & & 0 \\ & \tau_2 & \\ 0 & & \ddots \\ & & & \tau_m \end{bmatrix}$$

$$S_k = (s_{ii'}^k)_{n \times n} = \begin{bmatrix} s_{11}^k & s_{12}^k & \dots & s_{1i'}^k & \dots & s_{1n}^k \\ s_{21}^k & s_{22}^k & \dots & s_{2i'}^k & \dots & s_{2n}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ s_{i1}^k & s_{i2}^k & \dots & s_{ii'}^k & \dots & s_{in}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ s_{n1}^k & s_{n2}^k & \dots & s_{ni'}^k & \dots & s_{nn}^k \end{bmatrix}$$

$$\underline{t}^1 S_k = \begin{bmatrix} t_1^1 \\ t_1^1 \\ t_2^1 \\ \vdots \\ t_k^1 \\ \vdots \\ t_m^1 \end{bmatrix} \begin{bmatrix} s_{11}^k & s_{12}^k & \dots & s_{1i'}^k & \dots & s_{1n}^k \\ s_{21}^k & s_{22}^k & \dots & s_{2i'}^k & \dots & s_{2n}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ s_{i1}^k & s_{i2}^k & \dots & s_{ii'}^k & \dots & s_{in}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ s_{n1}^k & s_{n2}^k & \dots & s_{ni'}^k & \dots & s_{nn}^k \end{bmatrix}$$

avec $S_k = X_k^T X_k = (s_{ii'}^k)_{n \times n} = \left(\sum_{\ell=1}^n x_{i\ell}^k x_{\ell i'}^k \right)$

$$\underline{t}^1 s_k = t_1^1 s_1 + t_2^1 s_2 + \dots + t_k^1 s_k + \dots + t_m^1 s_m$$

avec $\underline{t}_k^1 s_k =$
$$\begin{bmatrix} t_k^1 s_{11}^k & \dots & t_k^1 s_{1i}^k & \dots & t_k^1 s_{1n}^k \\ t_k^1 s_{21}^k & \dots & t_k^1 s_{2i}^k & \dots & \\ \vdots & & \vdots & & \vdots \\ t_k^1 s_{n1}^k & \dots & t_k^1 s_{ni}^k & \dots & t_k^1 s_{nn}^k \end{bmatrix} = (t_k^1 s_{ii}^k)_{n \times n}$$

$$r_{n \times n}^1 = \sum_{k=1}^m \underline{t}_k^1 s_k = \left(\sum_{k=1}^m t_k^1 s_{ii}^k \right) = \left(t_1^1 s_{ii}^1 + t_2^1 s_{ii}^2 + \dots + t_m^1 s_{ii}^m \right)$$

avec $i = i' = 1, 2, \dots, n$

$$r^1 = \begin{bmatrix} \sum_{k=1}^m t_k^1 s_{11}^k & \dots & \sum_{k=1}^m t_k^1 s_{1n}^k \\ \vdots & & \vdots \\ \sum_{k=1}^m t_k^1 s_{n1}^k & \dots & \sum_{k=1}^m t_k^1 s_{nn}^k \end{bmatrix} = a a^T$$

Si $\{\rho_\ell; \underline{a}_\ell, \ell = 1, \dots, n\}$ sont les éléments propres de la matrice carrée $r^1: (n \times n)$, la factorisation canonique (ou diagonalisation) de r^1 peut s'écrire:

$$S_k = X_k^T X_k = \begin{bmatrix} x_{11}^k & \dots & x_{1\ell}^k & \dots & x_{1p_k}^k \\ \vdots & & \vdots & & \vdots \\ x_{i\ell}^k & \dots & x_{i\ell}^k & \dots & x_{ip_k}^k \\ \vdots & & \vdots & & \vdots \\ x_{n1}^k & \dots & x_{n\ell}^k & \dots & x_{np_k}^k \end{bmatrix} \begin{bmatrix} x_{11}^k & \dots & x_{1i'}^k & \dots & x_{1n}^k \\ \vdots & & \vdots & & \vdots \\ x_{\ell 1}^k & \dots & x_{\ell i'}^k & \dots & x_{\ell n}^k \\ \vdots & & \vdots & & \vdots \\ x_{p_k 1}^k & \dots & x_{p_k i'}^k & \dots & x_{p_k n}^k \end{bmatrix}$$

avec x_{12}^k de $X_k^T = x_{21}^k$ de X_k

ou x_{n1}^k de $X_k^T = x_{1n}^k$ de X_k

ou encore $x_{i\ell}^k$ de $X_k^T = x_{\ell i}^k$ de X_k

$$S_k = \left(\sum_{\ell=1}^m x_{i\ell}^k x_{\ell i'}^k \right) = \left(s_{ii'}^k \right)$$

avec $s_{ii'}^k = s_{i'i}^k$

$\Rightarrow$ La matrice réelle carrée $S_k: n \times n$ est symétrique, semi-définie positive; elle est définie positive après approximation selon la méthode des moindres carrés généralisés (transformation ⑨ dans le schéma d'ensemble des transformations).

Etant donné que:

- S_k : $n \times n$, d'élément $s_{ii}^k = \left(\sum_{\ell=1}^n x_{i\ell} x_{\ell i} \right)$ est une matrice réelle carrée symétrique, définie positive après approximation:

alors $\underline{t}^1 = \begin{bmatrix} t_1^1 \\ t_2^1 \\ \vdots \\ t_k^1 \\ \vdots \\ t_m^1 \end{bmatrix}_{m \times 1}$ est le 1^{er} vecteur propre de la matrice E réelle symétrique définie positive tel que:

$$(\underline{t}^1)^T \underline{t}^1 = \tau_1 \text{ avec } \tau_1 > 0;$$

$$\underline{r}^1 = \sum_{k=1}^m t_k^1 S_k$$

la matrice carrée réelle $\underline{r}^1$: $n \times n$ est symétrique définie positive.

Par conséquent, elle se factorise exactement.