

# STATISTIQUE ET ANALYSE DES DONNÉES

MARIE-ODILE PERNIN

MONIQUE PAGES

**Comparaison de deux méthodes : l'analyse des correspondances  
complète et l'analyse dissymétrique conditionnelle**

*Statistique et analyse des données*, tome 13, n° 3 (1988), p. 44-55.

[http://www.numdam.org/item?id=SAD\\_1988\\_\\_13\\_3\\_44\\_0](http://www.numdam.org/item?id=SAD_1988__13_3_44_0)

© Association pour la statistique et ses utilisations, 1988, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). 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/>

**Comparaison de deux méthodes :  
l'analyse des correspondances complète et  
l'analyse dissymétrique conditionnelle.**

Marie-Odile PERNIN et Monique PAGES  
Laboratoire de biométrie, URA CNRS 243, Université LYON I.  
43 bd. du 11 Novembre 1918, 69622 Villeurbanne Cedex

*Résumé : Lorsqu'un ensemble d'observations est structuré par deux variables qualitatives, l'espace vectoriel euclidien engendré par toutes les modalités de ces variables peut être partitionné différemment selon deux méthodes : l'Analyse des Correspondances Complète (ACC) et l'Analyse Dissymétrique Conditionnelle (ADC). Nous présentons dans cet article les relations théoriques existant entre les sous-espaces vectoriels qui résultent de ces deux décompositions, ainsi que les conséquences de ces relations pour l'interprétation concrète des résultats des analyses considérées.*

*Abstract : The Complete Correspondence Analysis (CCA) and Conditional Dissymmetrical Analysis (CDA) methods are based on a different partitioning of the euclidian space spanned by the categories of two categorical variables. In this paper, we present the theoretical relations existing between subspaces provided by these two decompositions, and the consequences of these relations relative to the concrete interpretation of results.*

**Mots clés :** analyse de variance à deux facteurs, analyse des correspondances complète, espaces euclidiens, projecteurs.

**Indices de classification STMA :** 08-110, 08-080, 06-110

## **1. Introduction.**

L'analyse des correspondances complète (ACC), proposée par Pontier et al. (4), est l'analyse canonique complète réalisée entre les fonctions indicatrices de deux variables qualitatives. Cette méthode s'applique notamment quand ces deux variables qualitatives sont des facteurs structurants de données quantitatives uni- ou multivariées. Cette situation correspond à un plan expérimental d'analyse de variance à deux facteurs contrôlés

Manuscrit reçu le 6 mai 1988

Révisé le 14 décembre 1988

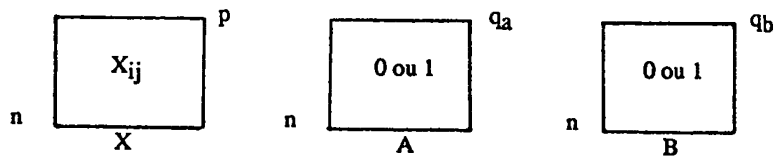
(MANOVA2) modèle à effets fixes : dans ce cas, d'une part Sabatier (5) et d'autre part Takeuchi et al. (7) montrent qu'une analyse, que nous appellerons Analyse Dissymétrique Conditionnelle (ADC), est réalisée entre les fonctions indicatrices des deux variables qualitatives.

L'objectif global de ces deux méthodes (ACC et ADC) est la mise en évidence de l'effet de tout ou partie de chacun des deux facteurs contrôlés sur la variabilité des données observées.

L'interprétation des deux méthodes (ACC et ADC) dans la structure d'espace euclidien, nous permet de comparer les décompositions de la variabilité observée obtenues par l'ADC puis par l'ACC et d'établir les relations qui existent entre ces décompositions.

## 2. Notations et proposition.

On peut considérer la juxtaposition des trois tableaux suivants comme une représentation schématique des données à analyser :



X est le tableau des données quantitatives, de dimension (n,p). Il contient les résultats de p mesures quantitatives (correspondant à p variables) faites sur n unités d'observation. On appelle  $\Omega$  l'ensemble des n unités d'observation.

A et B sont les tableaux disjonctifs complets associés aux deux variables qualitatives A et B, de dimensions respectivement (n,q<sub>a</sub>) et (n,q<sub>b</sub>), si q<sub>a</sub> (resp. q<sub>b</sub>) est le nombre de modalités de la variable A (resp. B). Ces deux tableaux contiennent les fonctions indicatrices des modalités de chacune des deux variables A et B.

\* Espaces euclidiens.

Chaque observation, notée  $\omega_i$ , est munie du poids  $p_i$  strictement positif, avec  $\sum_{i=1}^n p_i = 1$  (i=1 à n).  $[\Omega]$ , ensemble de fonctions réelles définies sur  $\Omega$ , est un espace vectoriel réel de dimension n. La pondération des observations permet de définir sur  $[\Omega]$  un produit scalaire f :

$$X, Y, \text{ éléments de } [\Omega], \quad f(X,Y) = \sum_{i=1}^n p_i X(\omega_i) Y(\omega_i) \quad (i=1 \text{ à } n)$$

On note  $X \perp_f Y$  (X f-orthogonal à Y) lorsque  $f(X,Y)=0$ , avec  $X \neq 0$  et  $Y \neq 0$

$[\Omega]$  est donc un espace euclidien dont une base  $f$ -orthonormale (notée  $\mathbb{B}$ ) est l'ensemble des  $n$  fonctions indicatrices des unités d'observation (on note  $\Omega_i$  la fonction indicatrice de  $\omega_i$ , avec  $\Omega_i(\omega_j) = \delta_{ij}$ ). On note  $D_n$  la matrice diagonale de terme général  $p_i$  ( $i=1$  à  $n$ ), matrice associée au produit scalaire  $f$  dans cette base.

On note de plus,  $[X]$ ,  $[A]$ ,  $[B]$  les sous-espaces vectoriels de  $[\Omega]$  engendrés respectivement par les colonnes des tableaux  $X$ ,  $A$ ,  $B$  ( $\dim [X] \leq p$ ,  $\dim [A] = q_a$ ,  $\dim [B] = q_b$ ), et  $[1]$  le sous-espace vectoriel de  $[\Omega]$  engendré par le vecteur noté  $1$  :

$$1 = (1,1,1,\dots,1) ; \quad [A] \supset [1] \text{ et } [B] \supset [1]$$

Enfin,  $[A+B]$  désigne le sous-espace vectoriel de  $[\Omega]$  engendré par les colonnes de  $A$  et de  $B$  réunies. On vérifie alors que  $\dim [A+B] \leq q_a + q_b - 1$ .

Soient  $[F]$  et  $[G]$  deux sous-espaces vectoriels quelconques de  $[\Omega]$  :

- . On note  $[F] \oplus [G]$ , la somme directe de  $[F]$  et  $[G]$  (avec  $[F] \cap [G] = \{0\}$ ).
- . On note  $[F] \oplus^\perp [G]$ , la somme directe  $f$ -orthogonale de  $[F]$  et  $[G]$  (avec  $[F] \perp [G]$ ).
- . On note  $[F]^\perp$ , le supplémentaire  $f$ -orthogonal de  $[F]$  dans  $[\Omega]$  :  $[\Omega] = [F] \oplus^\perp [F]^\perp$ .

. Dans la base  $\mathbb{B}$ ,  $P_F$  (resp.  $P_F^\perp = I_n - P_F$ , avec  $I_n$  matrice identité d'ordre  $n$ ) est la matrice du projecteur  $f$ -orthogonal sur  $[F]$  (resp. sur  $[F]^\perp$ ), et  $P_1$  celle du projecteur  $f$ -orthogonal sur  $[1]$ .

. On note  $P_F[G]$ , l'image de  $[G]$  par  $P_F$ . Dans le cas où il s'agit de  $P_F^\perp[G]$ , ce sous-espace vectoriel est noté  $[G/F]$ .

Proposition :

- i)  $[F+G] = [F] \oplus^\perp [G/F]$   
 $[F+G] = [G] \oplus^\perp [F/G]$
- ii) si  $[F] \perp [G]$  alors  $[F+G] = [F] \oplus^\perp [G]$
- iii) si  $[G] \supset [F]$  alors  $[F+G] = [G] \cap [F]^\perp \oplus^\perp [F]$   
 $[G/F] = [G] \cap [F]^\perp$

### 3. Décompositions de $[A+B]$ par l'ACC et l'ADC.

Le principe de chacune des deux méthodes (ACC et ADC) repose sur la décomposition de

l'espace  $[A+B]$  en une somme directe  $f$ -orthogonale de plusieurs sous-espaces. La projection de  $[X]$  sur ces sous-espaces permet la description de l'effet de chaque facteur contrôlé (A et/ou B), relativement aux caractéristiques de chaque sous-espace vectoriel de projection. Les carrés des normes des diverses projections des éléments de  $[X]$  sur ces sous-espaces correspondent aux différentes parts de variabilité observée dues à chaque facteur (A et/ou B). La somme des carrés de ces normes est égale à la variabilité totale observée (6) :

- dans le cas univarié, la variance totale de X est décomposée en somme de variances "partielles",
- dans le cas multivarié, la matrice de variances-covariances de X est décomposée en somme de matrices de variances-covariances "partielles".

La décomposition de  $[\Omega]$  en :  $[\Omega] = [A+B] \oplus [A+B]^\perp$  est utilisée de la façon suivante :

- La projection de  $[X]$  sur le sous-espace vectoriel  $[A+B]$  décrit la part de variabilité observée liée aux effets additifs des facteurs A et B.
- La projection de  $[X]$  sur  $[A+B]^\perp$  rend compte de la part de variabilité observée, soit comme variabilité résiduelle, soit comme la somme d'une part de variabilité observée relative à l'interaction des deux facteurs A et B et d'une part de variabilité résiduelle.

Dans un but de comparaison des deux méthodes, on ne s'intéresse qu'aux décompositions de  $[A+B]$ , le sous-espace vectoriel  $[A+B]^\perp$  étant lui-même redécomposé ou non.

### 3.1. Décomposition de $[A+B]$ par l'ADC.

L'ADC décompose le sous-espace vectoriel  $[A+B]$  en une somme  $f$ -orthogonale de trois sous-espaces (proposition i) :

$$[A+B] = [A] \oplus [B/A] = [1] \oplus [A/1] \oplus [B/A]$$

ou  $[A+B] = [B] \oplus [A/B] = [1] \oplus [B/1] \oplus [A/B]$

### 3.2. Décomposition de $[A+B]$ par l'ACC.

La décomposition complète de  $[A]$  (resp. de  $[B]$ ) est la suivante (1) :

$$[A] = ([A] \cap [B]) \oplus ([A] \cap [B])^\perp \oplus [A_2]$$

$$[B] = ([A] \cap [B]) \oplus ([B] \cap [A])^\perp \oplus [B_2]$$

$[A_2]$  (resp.  $[B_2]$ ) est le sous-espace vectoriel de  $[A]$  (resp.  $[B]$ ) dit "oblique" par rapport à  $[B]$  (resp. à  $[A]$ ).

Notons que l'analyse des correspondances complète (ACC) entre [A] et [B] fournit, les bases des sous-espaces  $[A] \cap [B]^\perp$  et  $[B] \cap [A]^\perp$  (sous-espaces propres associés aux valeurs propres nulles des matrices diagonalisées), en plus de celles données par l'analyse des correspondances ordinaire (4).

Ainsi, l'ACC fournit la décomposition complète de [A+B] en une somme directe *f*-orthogonale de cinq sous-espaces (et non pas de six, puisque [A<sub>2</sub>] et [B<sub>2</sub>] ne sont jamais *f*-orthogonaux) :

$$[A+B] = [1] \oplus ([1]^\perp \cap [A] \cap [B]) \oplus ([A_2] \oplus [B_2]) \oplus ([A] \cap [B]^\perp) \oplus ([B] \cap [A]^\perp)$$

Par ailleurs, on peut obtenir à partir de l'ACC, une décomposition du sous-espace [A+B] de [Ω] en six sous-espaces *f*-orthogonaux :

$$[A+B] = [1] \oplus ([1]^\perp \cap [A] \cap [B]) \oplus [A_2] \oplus [B'_2] \oplus ([A] \cap [B]^\perp) \oplus ([B] \cap [A]^\perp)$$

[B'<sub>2</sub>] est le supplémentaire *f*-orthogonal de [A<sub>2</sub>] dans ([A<sub>2</sub>] ⊕ [B<sub>2</sub>]) : les vecteurs *f*-orthogonaux qui complètent la base de [A<sub>2</sub>] dans ([A<sub>2</sub>] ⊕ [B<sub>2</sub>]) forment une base de [B'<sub>2</sub>].

On a donc :

$$[A_2] \oplus [B_2] = [A_2] \oplus [B'_2]$$

or, d'après la proposition i :

$$[A_2] \oplus [B_2] = [A_2] \oplus [B_2/A_2]$$

d'où :

$$[B'_2] = [B_2/A_2]$$

Notons qu'on peut obtenir la description de l'effet oblique de B (c'est-à-dire dépendant de (ou corrélé à) celui de A) en intervertissant les rôles de [A<sub>2</sub>] et [B<sub>2</sub>].

Nous montrons au paragraphe suivant que cette décomposition de [A+B] facilite la comparaison des décompositions issues de l'ACC et de l'ADC.

#### 4. Comparaison des décompositions de [A+B] issues de l'ACC et de l'ADC:

##### 4.1. Cas général.

Rappelons que les deux décompositions de [A+B] obtenues par l'ACC et l'ADC sont respectivement :

$$\text{ACC : } [A+B] = [1] \oplus ([1]^\perp \cap [A] \cap [B]) \oplus ([A_2] \oplus [B_2]) \oplus ([A] \cap [B]^\perp) \oplus ([B] \cap [A]^\perp)$$

$$\text{ADC : } [A+B] = [1] \oplus [A/1] \oplus [B/A]$$

$$\text{ou } [A+B] = [1] \oplus [B/1] \oplus [A/B]$$

La comparaison de ces deux décompositions permet d'établir les égalités suivantes :

$$\begin{aligned} [A/1] &= ([1]^\perp \cap [A] \cap [B]) \oplus^\perp [A_2] \oplus^\perp [A] \cap [B]^\perp \\ \text{et} \quad [B/A] \cap ([B_2] \oplus^\perp [B] \cap [A]^\perp) &= [B] \cap [A]^\perp \end{aligned}$$

Démonstration :

$$\begin{aligned} [A] \supset [1] &\Rightarrow [A/1] = [A] \cap [1]^\perp \quad (\text{proposition iii}) \\ \text{or} \quad [A] &= ([A] \cap [B]) \oplus^\perp ([A] \cap [B]^\perp) \oplus^\perp [A_2] \quad (\text{cf. §3.2}) \\ \text{donc} \quad [A] &= [1] \oplus^\perp ([1]^\perp \cap [A] \cap [B]) \oplus^\perp [A_2] \oplus^\perp ([A] \cap [B]^\perp) \\ \text{d'où} \quad [A/1] &= ([1]^\perp \cap [A] \cap [B]) \oplus^\perp [A_2] \oplus^\perp ([A] \cap [B]^\perp) \end{aligned}$$

ce qui conduit à :

$$\begin{aligned} \text{ACC} &: [A+B] = [A] \oplus ([B_2] \oplus^\perp [B] \cap [A]^\perp) \\ \text{ADC} &: [A+B] = [A] \oplus^\perp [B/A] \end{aligned}$$

Le choix du supplémentaire de  $[A]$  dans  $[A+B]$  n'est donc pas le même selon les deux méthodes : l'ACC n'utilise pas le supplémentaire  $f$ -orthogonal de  $[A]$ , afin d'obtenir des sous-espaces  $[A_2]$  et  $[B_2]$  "symétriques" (espaces obliques l'un par rapport à l'autre). Les relations entre ces deux supplémentaires sont alors :

$$\begin{aligned} [B/A] &= [A]^\perp \cap [A+B] & \text{d'où :} & \quad [B/A] \supset ([B] \cap [A]^\perp) \\ \text{et} \quad [B_2] \cap [A]^\perp &= \{0\} & (\text{par déf.}) & \\ [A]^\perp \supset [B/A] & & \text{d'où :} & \quad [B_2] \cap [B/A] = \{0\} \end{aligned}$$

c'est-à-dire que :  $[B/A] \cap ([B_2] \oplus^\perp [B] \cap [A]^\perp) = [B] \cap [A]^\perp$

De même en intervertissant les rôles de  $A$  et  $B$ , on montre que :

$$\begin{aligned} [B/1] &= ([1]^\perp \cap [A] \cap [B]) \oplus^\perp [B_2] \oplus^\perp ([B] \cap [A]^\perp) \\ \text{et} \quad [A/B] \cap ([A_2] \oplus^\perp [A] \cap [B]^\perp) &= [A] \cap [B]^\perp \end{aligned}$$

Ainsi, en raison de la non  $f$ -orthogonalité des sous-espaces  $[A_2]$  et  $[B_2]$ , on ne peut pas exprimer les différents sous-espaces issus de la décomposition de  $[A+B]$  selon l'ADC comme une somme de sous-espaces issus de la décomposition de  $[A+B]$  selon l'ACC.

Cependant, la décomposition dissymétrique de  $[A+B]$  fournie par l'ACC, proposée au paragraphe 3.2, permet une comparaison théorique plus satisfaisante des deux décompositions étudiées qui sont rappelons-le :

$$\text{ACC} : [A+B] = [1] \oplus^{\perp} ([1]^{\perp} \cap [A] \cap [B]) \oplus^{\perp} [A_2] \oplus^{\perp} [B_2] \oplus^{\perp} ([A] \cap [B]^{\perp}) \oplus^{\perp} ([B] \cap [A]^{\perp})$$

$$\text{ADC} : [A+B] = [1] \oplus^{\perp} [A/1] \oplus^{\perp} [B/A]$$

ou

$$\text{ACC} : [A+B] = [1] \oplus^{\perp} ([1]^{\perp} \cap [A] \cap [B]) \oplus^{\perp} [A'_2] \oplus^{\perp} [B_2] \oplus^{\perp} ([A] \cap [B]^{\perp}) \oplus^{\perp} ([B] \cap [A]^{\perp})$$

$$\text{ADC} : [A+B] = [1] \oplus^{\perp} [B/1] \oplus^{\perp} [A/B]$$

En effet, comme précédemment :  $[A/1] = ([1]^{\perp} \cap [A] \cap [B]) \oplus^{\perp} [A_2] \oplus^{\perp} ([A] \cap [B]^{\perp})$

c'est-à-dire que :  $\text{ACC} : [A+B] = ([1] \oplus^{\perp} [A/1]) \oplus^{\perp} ([B] \cap [A]^{\perp} \oplus^{\perp} [B'_2])$

$$\text{ADC} : [A+B] = ([1] \oplus^{\perp} [A/1]) \oplus^{\perp} [B/A]$$

mais cette fois,  $([B] \cap [A]^{\perp} \oplus^{\perp} [B'_2])$  est le supplémentaire  $f$ -orthogonal de  $[A]$  dans  $[A+B]$ . Ainsi, par unicité du supplémentaire  $f$ -orthogonal d'un sous-espace dans l'espace de référence :

$$[B/A] = ([B] \cap [A]^{\perp}) \oplus^{\perp} [B'_2]$$

Dans ce cas, la comparaison entre les deux décompositions de  $[A+B]$ , issues de l'ACC et de l'ADC, conduit aux égalités suivantes :

(cas 1)  $[A/1] = ([1]^{\perp} \cap [A] \cap [B]) \oplus^{\perp} [A_2] \oplus^{\perp} ([A] \cap [B]^{\perp})$   
 $[B/A] = [B'_2] \oplus^{\perp} ([B] \cap [A]^{\perp})$

et

(cas 2)  $[B/1] = ([1]^{\perp} \cap [A] \cap [B]) \oplus^{\perp} [B_2] \oplus^{\perp} ([B] \cap [A]^{\perp})$   
 $[A/B] = [A'_2] \oplus^{\perp} ([A] \cap [B]^{\perp})$

#### Exemple d'application numérique :

Dans cet exemple, traité par Dagnélie (2, §15.4.4.), l'objectif est de comparer trois types de sondes pédologiques dans deux types de sols afin de mettre en évidence l'effet du facteur sonde.

Tableau X des 22 valeurs observées (en mg de  $P_2O_5$  pour 100g de terre sèche) :

|       | sonde 1 | sonde 2 | sonde 3 |
|-------|---------|---------|---------|
| sol 1 | 43      | 41      | 42      |
|       | 45      | 42      | 44      |
|       | 46      | 44      | 46      |
|       | 53      |         |         |
| sol 2 | 40      | 35      | 37      |
|       | 40      | 37      | 39      |
|       | 40      | 40      | 40      |
|       | 43      | 40      | 40      |



A et B sont les tableaux disjonctifs complets associés respectivement aux variables qualitatives sol et sonde.

L'ADC met en évidence une part de variabilité due au facteur sonde sans le facteur sol ( $[B/A]$ ) de 47,8, tandis que cette part de variabilité mise en évidence par l'ACC ( $[B] \cap [A]^\perp$ ) est de 5,8. La différence (47,8 - 5,8) est présentée par l'ADC comme une part de variabilité due au facteur sonde, mais interdépendante des facteurs sol et sonde ( $[B]_2$ ). L'effet spécifique du facteur sonde n'est significatif, au risque de première espèce de 5%, que dans le cas de l'ADC (test de Fisher).

#### 4.2. Cas particulier du plan orthogonal.

Nous considérons la condition C suivante :  $[A/1] \perp [B/1]$ , sous laquelle le plan est dit "orthogonal" (ou complet) au sens de l'analyse de variance.

En effet :

$$\begin{aligned} \text{puisque (cf. 4.1)} \quad & [A/1] = [A] \cap [1]^\perp \\ \text{et} \quad & [B/1] = [B] \cap [1]^\perp \\ \text{sous la condition C :} \quad & ([A] \cap [1]^\perp) \perp ([B] \cap [1]^\perp) \\ \text{ce qui conduit à :} \quad & P_A P_B = P_B P_A = P_I \end{aligned}$$

c'est dire que la composition des deux projections de matrices  $P_A$  et  $P_B$  est équivalente à la projection sur  $[A \cap B] = [1]$ . Si de plus  $D_n = 1/n I_n$ , alors les effectifs de chaque classe de la partition de  $\Omega$ , induite par les deux variables qualitatives considérées simultanément, sont proportionnels ou égaux.

Ainsi, dans le cas d'un plan orthogonal, chacune des deux méthodes fournit une seule décomposition de  $[A+B]$  (proposition ii) :

$$\begin{aligned} \text{ACC :} \quad & [A+B] = [1] \oplus ([A] \cap [B]^\perp) \oplus ([B] \cap [A]^\perp) \\ \text{ADC :} \quad & [A+B] = [1] \oplus [A/1] \oplus [B/1] \end{aligned}$$

$$\text{avec} \quad [A/1] = [A] \cap [B]^\perp \quad \text{et} \quad [B/1] = [B] \cap [A]^\perp$$

Ces deux décompositions sont équivalentes.

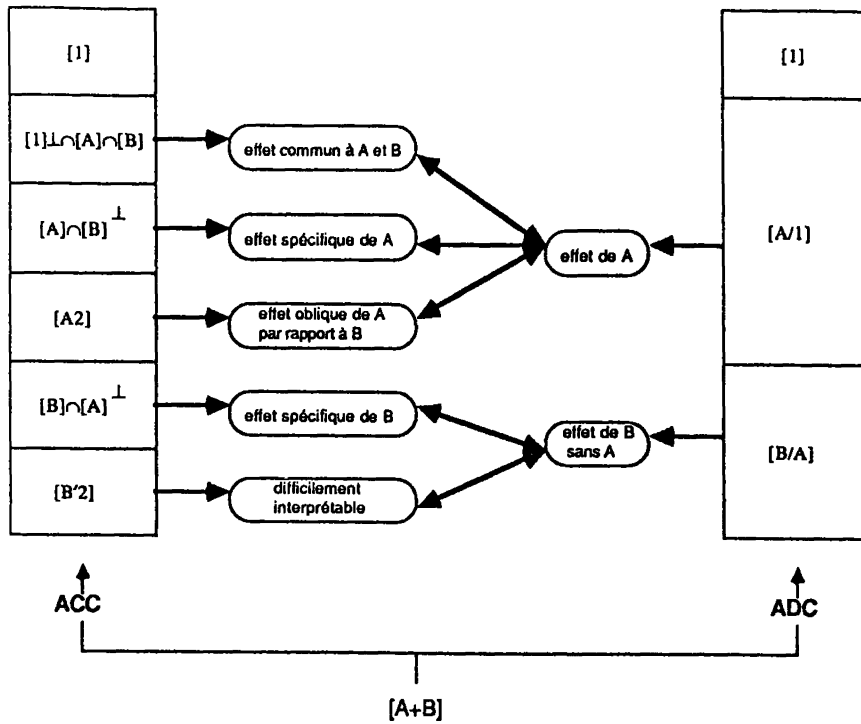


Schéma récapitulatif des relations entre les sous-espaces issus de la décomposition de  $[A+B]$  par l'ACC et par l'ADC

**5. Conclusion.**

Les deux décompositions de  $[A+B]$  obtenues par l'ACC (modifiée cf. §3.2) et l'ADC sont :

(cas 1) :

$$\text{ACC} : [A+B] = [1] \oplus ([1]^\perp \cap [A] \cap [B]) \oplus [A_2] \oplus [B'_2] \oplus ([A] \cap [B]^\perp) \oplus ([B] \cap [A]^\perp)$$

$$\text{ADC} : [A+B] = [1] \oplus [A/1] \oplus [B/A]$$

ou (cas 2) :

$$\text{ACC} : [A+B] = [1] \oplus ([1]^\perp \cap [A] \cap [B]) \oplus [A'_2] \oplus [B_2] \oplus ([A] \cap [B]^\perp) \oplus ([B] \cap [A]^\perp)$$

$$\text{ADC} : [A+B] = [1] \oplus [B/1] \oplus [A/B]$$

La projection de [X] sur les sous-espaces issus de la décomposition de [A+B] fournie par l'ACC, permet de distinguer (dans le cas 1) :

- a) - la part de variabilité observée due à un effet commun des deux facteurs A et B (obtenu par la projection de [X] sur  $[1]^\perp \cap [A] \cap [B]$ ),  
- la part de variabilité observée due à un effet du facteur A indépendant, au sens de la *f*-orthogonalité, de celui du facteur B, effet dit "spécifique" de A (obtenu par la projection de [X] sur  $[A] \cap [B]^\perp$ ),  
- la part de variabilité observée due à un effet du facteur A dépendant de (ou corrélé à) celui de B (obtenu par la projection de [X] sur l'oblique  $[A_2]$ ),

alors que la projection de [X] sur les sous-espaces issus de la décomposition de [A+B] fournie par l'ADC, par exemple dans MANOVA2, regroupe ces différentes parts de variabilité en une seule, dite "effet global de A" (obtenu par la projection de [X] sur  $[A/1]$ ).

- b) - la part de variabilité due à un effet "spécifique" du facteur B c'est-à-dire l'effet de B indépendant, au sens de la *f*-orthogonalité, de celui de A (obtenu par la projection de [X] sur  $[B] \cap [A]^\perp$ ).  
- la part de variabilité correspondant à la projection de [X] sur  $[B'_2]$ . Cette dernière est *a priori* difficilement interprétable puisque  $[B'_2]$  n'est pas le sous-espace "oblique" de [B] par rapport à [A], mais une projection de ce sous-espace sur  $[A_2]^\perp$ .

alors que dans MANOVA2, la projection de [X] sur  $[B/A]$  (issu de la décomposition de [A+B] fournie par l'ADC) regroupe ces deux parts de variabilité en une seule qui décrit l'effet du facteur B quand toute variation expliquée par le facteur A est éliminée.

[Dans le cas 2, les rôles de A et B sont intervertis].

Ainsi, les effets dûs seulement à chacun des facteurs, A ou B, sont différents selon qu'ils sont mis en évidence à l'aide de l'ACC ou de l'ADC : on peut penser que l'effet de chaque facteur mis en évidence par l'ACC ou l'ADC, n'est spécifique du facteur considéré (au moins au sens de la *f*-orthogonalité des sous-espaces) que dans le cas de l'ACC.

Notons que si A et B ne sont pas des variables qualitatives, mais des groupes de variables qualitatives ou même des variables quantitatives, rien ne change fondamentalement pour l'ADC (5).

Dans le cadre de l'utilisation pratique de l'ACC et de l'ADC, la bibliographie montre que les projections de  $[X]$  sur les sous-espaces vectoriels issus des décompositions de  $[A+B]$  sont utilisées différemment selon le but recherché :

- En statistique, des tests inférentiels permettent de comparer les différentes parts de variabilité (6) ;
- En analyse des données, on réalise par exemple une Analyse en Composantes Principales (ACP) des parts de  $[X]$  projetées sur les différents sous-espaces vectoriels ; les projections de  $[X]$  sur les sous-espaces vectoriels du type  $[B] \cap [A]^\perp$  sont décrites dans Pemin (3) (dans le cadre de la méthode LONGI). Les projections de  $[X]$  sur les sous-espaces vectoriels du type  $[B/A]$  sont étudiées par Sabatier (5) comme ACP sur Variables Instrumentales (ACPVI). Dans le cas où  $X$  est un tableau de contingence, on peut faire les mêmes analyses (ACC ou ADC) mais  $D_n$  est la matrice diagonale des fréquences conditionnelles déduites. Ainsi, dans le cadre de l'Analyse des Correspondances on obtient des résultats intéressants par projection d'une table de contingence sur les sous-espaces vectoriels du type  $[B] \cap [A]^\perp$  (10), mais également sur les sous-espaces du type  $[B/A]$  (8-9). Lorsque le plan est orthogonal, les décompositions de  $[A+B]$  en deux sous-espaces vectoriels, par les deux méthodes (ACC et ADC), sont identiques ; les ACP des projections de  $[X]$  sur chacun d'eux (avec la métrique de Mahalanobis) sont les analyses discriminantes de  $X$  structuré par les variables qualitatives A et B.

#### Références :

- 1) CAILLIEZ F., PAGES J.P. (1976) : "Introduction à l'analyse des données". Smash, 616 pages.
- 2) DAGNELIE (1975) : "Théorie et méthodes statistiques". Vol. 2, Les Presses Agronomiques de Gembloux Ed., 463 p.
- 3) PERNIN M.O. (1986) : "Contribution à la méthodologie d'analyse de données longitudinales. Exemple de la croissance chez l'être humain (Auxologie)". Thèse de doctorat, Lyon I, 265 pages.
- 4) PONTIER J., JOLICOEUR P., PERNIN M.O. (1987) : "Analyse canonique complète". Statistique et Analyse des Données (S.A.D.), 12, 124-148.

5) SABATIER R. (1987) : "Méthodes factorielles en analyse des données : approximations et prise en compte de variables concomitantes". Thèse d'état, U.S.T.L., 242 p.

6) SEARLE S.R. (1971) : "Linear Models", John Wiley and sons, London, 532 p.

7) TAKEUCHI K., YANAI H., MUKHERJEE B.N. (1982) : "The foundations of multivariate analysis"; Wiley Eastern Delhi, 458 p.

8) TER BRAAK C.F.J. (1987) : "Unimodal models to relate species to environment". Agricultural Mathematics Group, Box 100, NL-6700 AC Wageningen, The Netherlands, p. 83-89.

9) YANAI H. (1986) : "Some generalizations of correspondence analysis in terms of projection operators". In Data Analysis and Informatics, Vol. 4, E. Diday et al. eds., Elsevier North-Holland, 193-207.

10) YOCCOZ N., CHESSEL D. (1988) : "Ordination sous contraintes de relevés d'avifaune : élimination d'effets dans un plan d'observations à deux facteurs (Ecologie générale)". C. R. Acad. Sci. Paris, t. 307, Série III, p. 189-194, 1988.