

REVUE DE STATISTIQUE APPLIQUÉE

PH. CASIN

L'analyse discriminante de tableaux évolutifs

Revue de statistique appliquée, tome 43, n° 3 (1995), p. 73-91

http://www.numdam.org/item?id=RSA_1995__43_3_73_0

© Société française de statistique, 1995, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

L'ANALYSE DISCRIMINANTE DE TABLEAUX ÉVOLUTIFS

Ph. Casin

*Faculté de Droit et d'Economie
Ile du Saulcy
57 005 METZ Cedex 1*

RÉSUMÉ

A partir du critère de généralisation de l'analyse canonique de CARROLL, une autre technique d'analyse de tableaux ternaires est obtenue en imposant des contraintes d'orthogonalisation dans chaque espace.

Mots-clés : *Analyse canonique généralisée, analyse discriminante, tableaux évolutifs.*

ABSTRACT

ADT is a method of analyzing 3-way data tables using CARROLL's criteria of generalized canonical analysis, under a constraint of orthogonalisation in each subspace. ADTE is the application of ADT at indexed data tables.

Keywords : *generalized canonical analysis, discriminant analysis, multivariate time series*

Introduction

L'analyse canonique généralisée, au sens de CARROLL (CARROLL (68)) présente un intérêt théorique certain, puisque les méthodes usuelles d'analyse des données en sont des cas particuliers.

L'interprétation des résultats de l'AGG est cependant délicate. L'analyse discriminante de tableaux (ADT) présentée ici est une variante de la méthode de CARROLL obtenue en modifiant la contrainte d'orthogonalisation; les résultats deviennent alors plus lisibles.

Un exemple d'application – l'évolution de 11 pays européens depuis 1965 – nous permettra d'illustrer l'ADT dans un de ses cas particuliers, l'analyse discriminante de tableaux évolutifs (ADTE).

1. Les données et leur représentation

1.1. Les données

On dispose de K tableau X_k ; chaque tableau X_k $k = 1, \dots, K$ décrit les valeurs prises par un ensemble de n mêmes individus – les lignes du tableau – pour m_k variables – les colonnes du tableau.

Les données décrites par le tableau X_k sont centrées pour chaque variable. On suppose de surcroît, sans perdre en généralité, que le poids de chaque individu, est identique et égal à $1/n$ et d'autre par que les m_k variables décrites par un tableau X_k sont linéairement indépendantes.

1.2. L'espace R^n

Chaque variable, pour chaque tableau, étant définie par les valeurs qu'elle prend pour les n individus peut être représentée par un vecteur de R^n .

La matrice diagonale des poids des individus D_p confère à R^n une structure d'espace vectoriel euclidien; le produit scalaire de deux vecteurs-variable est égal à leur covariance.

Dans R^n , les colonnes de chacun des tableaux X_k , $k = 1, \dots, K$ engendrent un espace vectoriel W_k , $k = 1, \dots, K$ de dimension égale à m_k ; la somme des espaces W_k , $k = 1, \dots, K$ est un espace de dimension au plus égale à $m = \sum_{k=1}^K m_k$, dont un système générateur est donné par les colonnes de $X = [X_1, \dots, X_k, \dots, X_K]$.

1.3. L'espace R^m

Soit X^* le tableau de dimensions nK et m suivant :

$$X^* = \begin{bmatrix} X_1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & X_k & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & X_K \end{bmatrix}$$

X décrit un nuage de points N dans l'espace des individus R^m ; si on suppose que R^m est muni de la métrique euclidienne usuelle, X^* est obtenu en considérant les projections orthogonales de N sur chacun des sous-espaces engendrés par les tableaux X_k , $k = 1, \dots, K$.

La matrice D_p^* des poids des individus est :

$$D_p^* = \frac{1}{nK} Id_{nK}$$

où Id_{nK} désigne la matrice unité d'ordre nK .

2. Rappels sur l'analyse canonique généralisée

2.1. Le critère à maximiser

Dans le cas de tableaux évolutifs, l'analyse canonique généralisée (CARROLL (68), SAPORTA (75) et (81), CASIN et TURLLOT (86)) se propose de rechercher dans une première étape une variable Z^1 et K variables synthétiques $Z_k^1 = X_k b_k^1$, résumant chaque tableau X_k ($k = 1, \dots, K$) b_k^1 étant un vecteur-colonne comportant m_k lignes de telle manière que :

$$\sum_{k=1}^K R^2(Z^1, Z_k^1)$$

soit maximal.

Z^1 est la première composante canonique.

Cette recherche se poursuit sous une contrainte d'orthogonalisation portant sur les variables canoniques Z^j ; à l'étape j , on détermine Z^j et les Z_k^j de telle façon que :

$$\sum_{k=1}^K R^2(Z^j, Z_k^j)$$

soit maximal, sous les contraintes :

$$\forall r < j, (Z^r)' D_p Z^j = 0$$

Les variables Z^j étant définies à une constante multiplicative près, on convient que :

$$\forall j = 1, \dots, m, (Z^j)' D_p Z^j = 1$$

2.2. L'ACG est une analyse en composantes principales

Dans R^n , le coefficient de détermination entre Z^j et Z_k^j est maximal lorsque Z_k^j est la projection D_p orthogonale de Z^j sur W_k .

Si on note P_k le D_p -projecteur sur l'espace W_k :

$$P_k = X_k ((X_k)' D_p X_k)^{-1} (X_k)' D_p$$

alors,

$$\sum_{k=1}^K R^2(Z^j, Z_k^j) = (Z^j)' \left(\sum_{k=1}^T P_k \right) Z^j$$

a une valeur maximale lorsque Z^j est le vecteur propre associé à la plus grande valeur propre de :

$$\sum_{k=1}^K P_k = X(K(X^*)'D_p^*X^*)^{-1}X'D_p$$

Cette valeur maximale est égale à cette plus grande valeur propre.

Par conséquent, l'AGG est l'ACP dans la métrique $(K(X^*)'D_p^*X^*)^{-1}$ du tableau X .

Les facteurs successifs B^j , $j = 1, \dots, m$ sont de la forme :

$$B^j = \begin{bmatrix} b_1^j \\ b_k^j \\ b_K^j \end{bmatrix}$$

2.3. L'ACG est une analyse factorielle discriminante

Dans R^m , l'ACG est aussi l'analyse factorielle discriminante du tableau X^* , la partition des individus en groupes étant donnée par le tableau G à nK lignes et n colonnes suivant :

$$G^\circ = \begin{bmatrix} Id_n \\ Id_n \\ Id_n \end{bmatrix}$$

Id_n désignant la matrice identité d'ordre n .

En effet, l'analyse factorielle discriminante est une ACP particulière, celle du tableau des centres de gravité des groupes effectuée dans la métrique donnée par l'inverse de la matrice de variance totale (CAILLIEZ et PAGES (76), ROMEDER (73)).

Ici, le tableau des centres de gravité est le tableau :

$$(G^\circ)'D_p^*X^* = \frac{1}{K}X$$

et l'inverse de la matrice de variance totale est donnée par $((X^*)'D_p^*X^*)^{-1}$, et on retrouve le résultat du paragraphe précédent.

Les pouvoirs discriminants sont égaux aux valeurs propres de l'AGG divisées par K , c'est-à-dire à la somme des R^2 entre les Z^j et les Z_k^j divisée par K .

2.4. Remarques

D'autres généralisations de l'analyse canonique ont été proposées (HORST (61), KETTENRING (71), Mc KEON (66), PONTIER et NORMAND (92)); la généralisation de CARROLL est la plus connue en France, car elle admet comme cas particulier l'analyse factorielle des correspondances multiples.

Mais il n'est pas certain (ESCOFIER, PAGES (88)) que les composantes Z^j soient bien liées linéairement aux variables de départ.

D'autre part, comme le fond remarquer PONTIER et NORMAND, (PONTIER et NORMAND (92)), l'analyse de CARROLL ne permet pas d'obtenir une base orthogonale de chacun des espaces W_k , $k = 1, \dots, K$; obtenant m composantes canoniques pour chaque espace de dimension m_k , certaines d'entre elles sont forcément ou nulles ou combinaison linéaire des autres.

L'analyse de CARROLL présente cependant deux propriétés intéressantes et utiles à l'interprétation des résultats, la maximisation d'une somme de R^2 et le caractère discriminant des axes canoniques; l'analyse discriminante de tableaux, proposée ci après, conserve ces deux propriétés tout en déterminant une base de chacun des sous-espaces constituée de variables deux à deux non corrélées.

3. L'analyse discriminante de tableaux (ADT)

3.1. Le problème

Il s'agit de reprendre le critère de CARROLL en imposant aux variables canoniques de chaque espace d'être D_p -orthogonale entre elles; à l'étape un, comme en ACG, on cherche une variable générale Z^1 et on résume chaque tableau X_k , $k = 1, \dots, K$, par une variable synthétique

$$Z_k^1 = X_k b_k^1$$

b_k^1 étant un vecteur colonne comportant m_k lignes de telle manière que :

$$\sum_{k=1}^K R^2(Z^1, Z_k^1)$$

soit maximal.

Cette recherche se poursuit sous une contrainte d'orthogonalisation qui porte ici sur les variables canoniques Z_k^1 ; à l'étape j , on détermine Z^j et les Z_k^j de telle façon que :

$$\sum_{k=1}^K R^2(Z^j, Z_k^j)$$

soit maximal, sous les contraintes :

$$\forall k, \quad \forall r < j, \quad (Z_k^r)' D_p Z_k^j = 0$$

On ne cherche donc plus à déterminer une base D_p -orthogonale de R^n , mais on recherche une base D_p -orthogonale pour chacun des sous-espaces W_k .

3.2. La solution

A l'étape un, l'ADT se confond avec l'ACG et Z^1 s'obtient en diagonalisant la somme des D_p -projecteurs sur les espaces W_k , $k = 1, \dots, K$.

Les Z_k^1 sont obtenus en projetant Z^1 sur W_k , $k = 1, \dots, K$.

Notons W_k^2 le sous-espace de W_k orthogonal à Z_k^1 ; W_k^2 est un espace de dimension $m_k - 1$ (ou m_k si la projection de Z_k^1 sur W_k est le vecteur nul), dont un système générateur est donné par les résidus des régressions des m_k variables de départ par la variable canonique Z_k^1 .

A l'étape deux, les K contraintes d'orthogonalisation imposent que les K composantes canoniques appartiennent aux espaces W_k^2 ; par conséquent, il s'agit de la première étape d'une ACG sur ces espaces.

Z^2 est le vecteur propre associé à la plus grande valeur propre de la somme des projecteurs sur les espaces W_k^2 , les composantes canoniques Z_k^2 sont les projections de Z^2 sur W_k^2 .

Par récurrence, on peut ainsi déterminer m_k composantes canoniques pour chacun des K espaces.

Dans la suite de l'exposé, on notera W_k^{j+1} le sous-espace de W_k D_p -orthogonal aux j premières composantes canoniques Z_k^1, \dots, Z_k^j et P_k^{j+1} le D_p projecteur sur W_k^{j+1} .

Cet espace W_k^{j+1} est engendré par les colonnes d'un tableau que l'on notera X_k^{j+1} ; le tableau X_k^{j+1} est obtenu à partir des résidus des régressions des m_k variables de X_k sur les j premières composantes canoniques de l'espace W_k .

$$Z^{j+1} \text{ est le vecteur propre de } \sum_{k=1}^K P_k^{j+1}$$

associé à sa valeur propre la plus élevée, notée v_{j+1} .

L'analyse discriminante de tableaux comporte donc au moins $\sup(m_k)$ étapes (et exactement $\sup(m_k)$ si aucune D_p orthogonalité entre des sous-espaces n'est décelée) et se termine lorsque les K espaces sont décrits.

A l'étape j , elle s'identifie avec la première étape d'une ACG portant sur les espaces W_k^j ; elle est donc la première étape d'une analyse discriminante du tableau :

$$X^{*j} = \begin{bmatrix} X_1^j & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & X_k^j & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & X_{K_j}^j \end{bmatrix}$$

contre le tableau G° défini au paragraphe 2.3, K_j étant le nombre d'espaces W_k^j ne se réduisant pas au vecteur nul.

Les pouvoirs discriminants d_j correspondants sont égaux aux valeurs propres v_j divisées par K_j .

Comme l'AGG ou la méthode LONGI (PERNIN (86), PONTIER et PERNIN (89), PONTIER, DUFOUR et NORMAND (90)), la présentation de la méthode en termes de discrimination permet une extension aux cas de données manquantes, ou aux cas où l'on dispose de plusieurs mesures pour une même série d'observations.

3.3. Les propriétés de l'analyse discriminante de tableaux

Propriété 1 : l'ADT généralise l'analyse canonique.

Preuve : En effet, les contraintes d'orthogonalisation sont les mêmes que celles de l'analyse canonique simple et chercher Z^j , Z_1^j et Z_2^j telles que $R^2(Z^j, Z_1^j) + R^2(Z^j, Z_2^j)$ soit maximal revient à maximiser $R^2(Z^j, Z_1^j)$ (SAPORTA. (75), par exemple).

Remarques

Contrairement à ce qui se passe en AGG, on obtient ici exactement le même nombre de composantes canoniques qu'en effectuant une analyse canonique simple.

Et, puisque l'analyse canonique simple est un cas particulier de l'ADT, l'analyse des correspondances multiples et l'analyse en composantes principales normée sont des cas particuliers de l'ADT (CASIN et TURLOT (86)).

Propriété 2 : La variable Z^j est la première composante principale, au sens de l'ACP normée, des variables canoniques Z_k^j .

Preuve : Les variables canoniques Z_k^j étant données, Z^j est la variable qui rend maximale $\sum_{k=1}^K R^2(Q, Z_k^j)$ donc par définition la première composante principale des Z_k^j .

Remarque : En AGG, Z^j est composante principale, au sens de l'ACP normée des Z_k^j , mais pas forcément la première composante principale.

Propriété 3 : A une constante multiplicative près, la variable discriminante Z^j est la somme des variables canoniques Z_k^j .

Preuve :
$$\sum_{k=1}^K Z_k^j = \sum_{k=1}^K P_k^j Z^j = v_j Z^j$$

v_j , défini au paragraphe précédent, étant la plus grande valeur propre de la somme des P_k^j .

Remarque : cette propriété est également vérifiée en ACG.

Propriété 4 : Si $r > j$, alors Z_k^r est D_p -orthogonale à Z^j .

Preuve : Par construction, Z_k^r est D_p -orthogonal à Z_k^j ; d'autre part, Z_k^j étant la projection de Z^j sur W_k^j , on a : $Z^j = Z_k^j + e$, avec e D_p -orthogonal à tout vecteur de W_k^j , et donc en particulier D_p -orthogonal à Z_k^r .

Donc Z^j et Z_k^r sont D_p -orthogonaux.

Propriété 5 : Les variables discriminantes Z^j sont deux à deux D_p -orthogonales.

Preuve : Considérons deux variables discriminantes Z^j et Z^r , et supposons que $r > j$.

D'après le propriété 3 :

$$\sum_{k=1}^K Z_k^r = v_r Z^r$$

D'après le propriété 4 :

$$k = 1, \dots, K \quad (Z_k^r)' D_p Z^j = 0$$

ce qui prouve le résultat annoncé.

Propriété 6 : v_2 est inférieure ou égale à la deuxième valeur propre de l'ACG portant sur les mêmes tableaux.

preuve : La première composante principale de l'ACG est la même que la première composante de l'ADT.

La deuxième composante principale de l'ACG est obtenue sous contrainte d'orthogonalisation avec la première; en ADT, la contrainte est multiple et porte sur chacune des K variables canoniques; d'après la propriété 5, si ces K contraintes sont vérifiées, la contrainte d'orthogonalisation avec la première composante principale est remplie.

A l'étape 2, l'ADT est donc une ACG avec des contraintes supplémentaires ce qui prouve le résultat annoncé.

Propriété 7 : Les valeurs propres sont comprises entre 1 et K .

Preuve : les valeurs propres sont égales à

$$v_j = \sum_{k=1}^K R^2(Z^j, Z_k^j)$$

Si un vecteur Q est commun aux K espaces W_k , alors Q , Z^j et Z_k^j sont colinéaires et v_j est égal à K .

D'autre part, dans le cas où les espaces W_k^j sont D_p -orthogonaux, la matrice de corrélation obtenue en prenant dans chacun de ces espaces un vecteur quelconque est la matrice identité. Par conséquent, d'après la propriété 2, la valeur propre est égale à 1.

La réciproque est vraie : si, à l'étape j , la valeur propre est égale à 1, cela signifie que toutes les valeurs propres de la matrice de corrélations sont égales à 1 et donc que les espaces W_k^j sont orthogonaux.

Remarques

Soit R le nombre d'étapes que comporte l'ADT :

$$D_1 = \sum_{j=1}^R (v_j - 1)^2 \quad \text{ou} \quad D_2 = \sum_{j=1}^R (v_j - 1)$$

sont les indicateurs de la proximité globale entre les K espaces.

Si les K espaces sont confondus, R est leur dimension commune et $D_2 = R(K-1)$ et $D_1 = R(K-1)^2$.

Si les K espaces sont D_p -orthogonaux, toutes les valeurs propres sont égales à 1 et D_1 et D_2 sont nuls.

Dans le cas particulier où $K = 2$ (analyse canonique ordinaire), si on note r_i les S coefficients de corrélation canoniques correspondant aux composantes canoniques non-orthogonales à l'un des deux espaces :

$$D_1 = \sum_{i=1}^S r_i^2 \quad \text{et} \quad D_2 = \sum_{i=1}^S r_i$$

La définition d'un tel indicateur est délicate en ACG, car d'une part la somme des m valeurs propres est toujours égales à m , d'autre part les composantes canoniques de chaque espace ne sont pas linéairement indépendantes.

3.4. Les représentations graphiques et les aides à l'interprétation

La représentation des individus est du même type que celle obtenue en ACG (CASIN (85), CASIN et TURLLOT (86)) : pour chaque axe, on représente $X^* B^j$, donc les variables Z_k^j et les centres de gravité des groupes soit $d_j Z^j$ (d_j , rappelons-le, désignant le $j^{\text{ème}}$ pouvoir discriminant).

Si le nombre d'individus est élevé, les graphiques ne sont pas très lisibles; il est alors sans doute plus judicieux de représenter les individus dans chaque espace W_k et de comparer les images obtenues.

Pour faciliter l'interprétation des résultats, il est nécessaire comme en ACG de calculer l'homogénéité de chaque groupe pour chacun des axes discriminants : si un groupe a une variance élevée pour un axe donné, cela signifie que ce groupe a un comportement original pour le phénomène mis en évidence.

Les composantes canoniques Z_k^j étant deux à deux D_p -orthogonales, les variables de chaque espace W_k sont représentées par les corrélations avec ces composantes. On peut ainsi, d'un coup d'œil, juger de l'inertie des variables de départ avec les axes considérés.

3.5. L'analyse discriminante de tableaux évolutifs

Dans le cas de tableaux évolutifs on dispose d'observations portant sur O objets définis pour V variables à T instants différents.

Si O et T sont suffisamment élevés, deux types d'analyse sont possibles :

- une analyse dans laquelle les O objets sont les individus; on dispose alors de T tableaux
- une analyse dans laquelle les T instants constituent les individus; on dispose alors de O tableaux

Ces deux analyses constituent la méthode ADTE, analyse discriminante de tableaux évolutifs.

L'ADTE est assez proche de la méthode LONGI; la différence entre les deux méthodes tient à la manière dont sont construites les variables du type Z_k^j : celles-ci sont deux à deux D_p -orthogonales et permettent une description complète de chaque sous-espace dans le cas de l'ADTE, alors que dans le cas de LONGI, pour une étape donnée, elles sont construites de la même façon pour les différents tableaux, c'est-à-dire que b_k^j ne dépend pas de k .

4. L'évolution de 11 pays depuis 1965

4.1. Les données

Les données (ÉCONOMIE EUROPÉENNE (93)) caractérisent la situation de 11 pays européens : Belgique (B), Danemark (D), Royaume-Uni (R), Italie (I), France (F), Irlande (E), Portugal (P), Luxembourg (L), Pays-bas (H), Espagne (S), Grèce (G).

Pour chacun de ces 11 pays, on dispose des variables suivantes :

- Taux de chômage (CHO)
- Taux de croissance du PIB (CRO)
- Taux d'inflation (INF)
- Part de la consommation privée dans le PIB (CSO)
- Taux d'investissement (INV)

pour chacune des années 1965, 1968, 1971, 1974, 1977, 1980, 1983, 1986, 1989 et 1992.

En ce qui concerne les représentations graphiques, l'intitulé des pays sera abrégé par les trois premières lettres pour les cercles de corrélations et les centres de

gravité des groupes; ces centres de gravité sont soulignés dans les graphiques. Les représentations des individus au cours du temps se font à l'aide de trois caractères; le premier caractérise le pays comme indiqué ci-dessus entre parenthèses, les deux derniers précisent l'année : ainsi, *S65* désigne l'Espagne en 1965.

Il s'agit plus ici d'illustrer une méthode que de traiter complètement un jeu de données, et c'est pourquoi nous limiterons l'interprétation des résultats de l'ADTE aux deux premières étapes de chacune des deux analyses.

En fait, comme pour les autres méthodes de description de tableaux ternaires, les résultats les plus intéressants de chaque analyse ne sont pas forcément ceux fournis par les premières étapes; une relation linéaire forte entre un nombre restreint de tableaux peut être décrite par un axe d'ordre élevé.

Les calculs peuvent être faits indifféremment par un logiciel d'analyse discriminante ou un logiciel d'ACG, moyennant la rédaction d'un programme effectuant les régressions à chaque étape.

4.2. L'analyse des 11 groupes-pays

L'ADT des 10 tableaux détermine 5 axes dont les pouvoirs discriminants sont : 0.931, 0.853, 0.738, 0.509, 0.342.

Pour les deux premiers axes, le tableau suivant indique la corrélation entre les variables Z^j ($j = 1, 2$) et les composantes Z_k^j ($k = 1, \dots, 10$) définies sur chacun des tableaux.

Tableau des corrélations

	Axe 1	Axe 2
65	.91	.79
68	.97	.92
71	.98	.92
74	.96	.95
77	.96	.87
80	.98	.89
83	.96	.96
86	.98	.98
89	.96	.96
92	.98	.96

L'analyse des corrélations avec les variables se fait grâce aux cercles de corrélations obtenus pour chaque tableau.

En ce qui concerne les individus, outre leur représentation pour les deux premières composantes principales, on utilise les variances des groupes pour les différents axes.

Cette variance des groupes a pour valeur moyenne la variance intraclasse, qui elle-même est égale à $1 - d_j$, d_j désignant le $j^{\text{ème}}$ pouvoir discriminant.

Variances des groupes

	Axe 1	Axe 2
Belgique	.049	.188
Danemark	.034	.119
Grèce	.084	.111
Espagne	.092	.278
France	.030	.158
Irlande	.115	.235
Italie	.030	.194
Luxembourg	.049	.066
Hollande	.038	.134
Portugal	.164	.008
Royaume-uni	.065	.005

Ces tableaux et ces graphiques suggèrent les interprétations suivantes :

– pour l’axe 1, de 1965 à 1992, et surtout jusqu’en 1980, on observe une très forte corrélation avec le taux de consommation; du point de vue des individus, l’axe 1 oppose d’une part, la Grèce, le Portugal, l’Espagne, pays où ce taux est très élevé, au Luxembourg et au Danemark, d’autre part, où le taux est moins élevé. Cette hiérarchie des taux est moins marquée dans les années récentes .

A partir de 77, l’axe 1 est assez corrélé à l’inflation; les pays à fort taux de consommation tout au long de la période d’observation sont ceux qui ces dernières années ont le plus mal maîtrisé l’inflation.

– pour l’axe 2, en 74 et après, la corrélation est forte avec le taux de chômage; avant 74, c’est le taux d’inflation qui est le plus corrélé à cet axe. Cet axe oppose l’Irlande et l’Espagne, pays à forte inflation avant 74 et à fort chômage après, à la Grèce, au Royaume-Uni et à la Belgique, qui connaissent dans les années récentes un taux de chômage plus faible et dont l’inflation était modérée avant 74.

Les variances résiduelles de l’Espagne et de l’Irlande sont élevées car leur comportement n’est pas constant au cours du temps : ainsi, en 77, le taux de chômage de l’Espagne est, relativement aux autres pays, moins élevé que les années suivantes et en 65, le taux d’inflation de l’Irlande est faible.

4.3. L’analyse des 10 groupes-instants

Les pouvoirs discriminants valent respectivement : 0.957, 0.878, 0.735, 0.536 et 0.327.

Comme pour l’analyse précédente, deux types de graphiques et deux tableaux sont utilisés pour interpréter les axes.

Tableau des corrélations

	Axe 1	Axe 2
Belgique	.99	.87
Danemark	.92	.93
Grèce	.97	.98
Espagne	.99	.94
France	.99	.96
Irlande	.99	.91
Italie	.97	.99
Luxembourg	.92	.94
Hollande	.99	.93
Portugal	.99	.92
Royaume-uni	.99	.93

Variances de groupes

	Axe 1	Axe 2
65	.024	.081
68	.044	.083
71	.030	.074
74	.006	.077
77	.070	.119
80	.024	.273
83	.031	.087
86	.105	.104
89	.031	.163
92	.063	.150

Pour tous les pays, à l'exception de la Grèce où le phénomène est moins marqué, l'axe 1 est celui du chômage : celui-ci augmente de période en période jusqu'au milieu des années 80, puis diminue légèrement après. La Grèce a un comportement particulier : le taux de chômage est plus faible dans les années 70 que dans les années 60.

L'axe 2 est très lié à l'inflation pour la plupart des pays : augmentant régulièrement de 65 à 71, très forte de 74 à 80, elle diminue beaucoup ensuite, surtout à partir de 86.

L'année 80 et les années 89 et 92 ont de fortes variances résiduelles : ces années-là, les 11 pays ont des taux d'inflation assez différents.

4.4. Comparaison avec les résultats de l'ACG

Pour chacune des deux analyses, à rang égal, les valeurs propres de l'ADTE sont inférieures (ou, bien entendu, égale pour la première valeur propre) aux valeurs de l'ACG; dans le premier cas, les pouvoirs discriminants de l'ACG valent

respectivement : 0.931, 0.854, 0.748, 0.571, 0.490, et dans le second cas 0.957, 0.879, 0.744, 0.602, 0.559.

Les résultats de l'ACG et de l'ADT sont peu différents aux premières étapes : en effet, lorsque comme ici, Z^j est très liée linéairement à toutes les variables Z_k^j , pour $k = 1, \dots, K$ les deux systèmes de contraintes sont très proches.

En effet, la détermination de Z^{j+1} se fait sous la contrainte de D_p orthogonalité à Z^j dans le cas de l'ACG ; à la limite si Z^j coïncide avec les Z_k^j $k = 1, \dots, K$, Z^{j+1} est D_p -orthogonale à Z_k^j et donc Z_k^{j+1} est D_p -orthogonale à Z_k^j .

Rappelons cependant que dans le cas de l'ADT les représentations graphiques sont plus complètes puisqu'elles se font pour chacun des tableaux.

5. Conclusion

L'ADT et l'ADTE apparaissent comme des méthodes intermédiaires entre l'ACG de K tableaux et l'ACP de chacun des tableaux.

En effet, la description de chaque tableau X_k est donnée par une ACP de métrique M_k , M_k reflétant les relations entre X_k et les autres tableaux X_j , $j \neq k$.

La comparaison entre ces K ACP se fait à l'aide des deux critères, le critère discriminant pour juger des proximités entre individus et le critère de CARROLL qui mesure la proximité globale du tableau pour l'axe considéré.

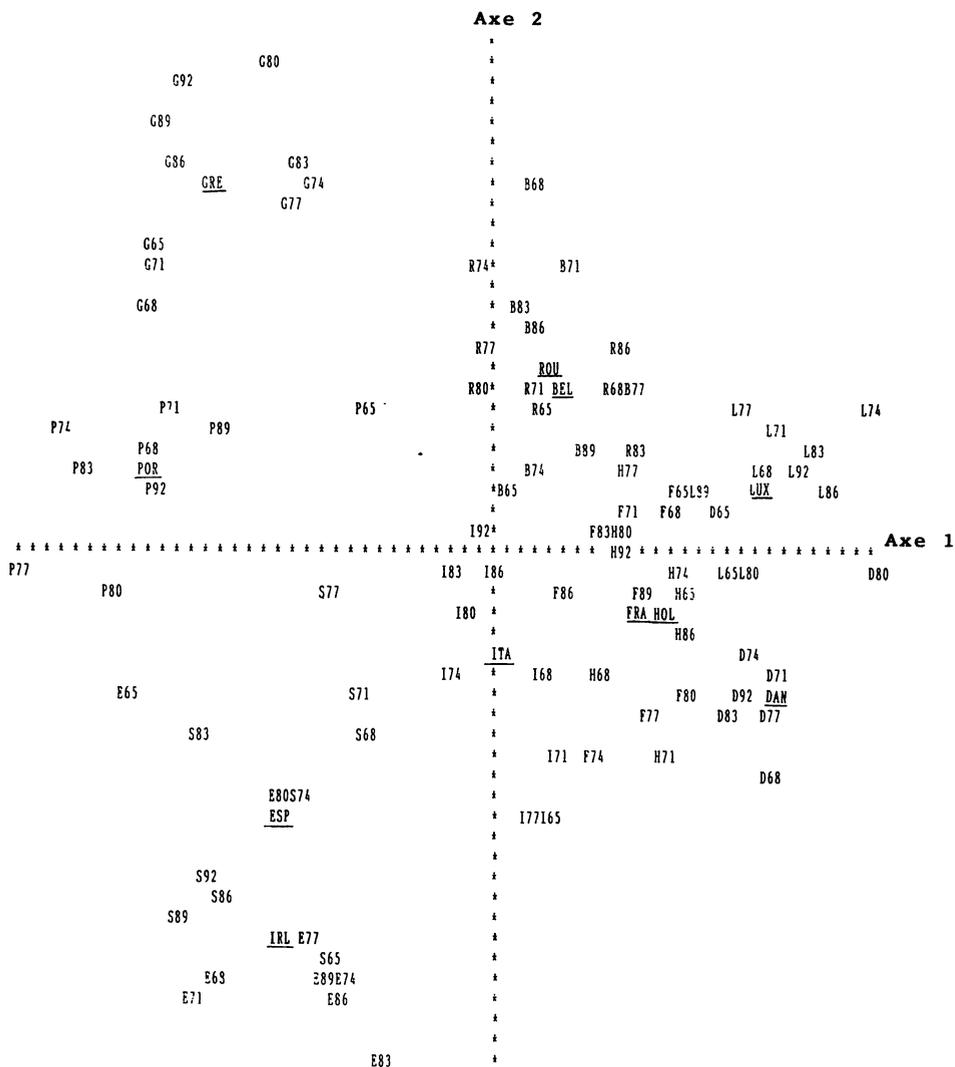
Ainsi, une seule technique permet une description des tableaux et une description des relations entre ces tableaux.

Bibliographie

- CAILLIEZ F. et PAGES J.P. – «Introduction à l'analyse des données», SMASH 1976.
- CARROLL J.D. – «A generalization of canonical correlation analysis to three or more sets of variables», Proc; 76th Conv. Amer. Psych. As., 1968
- CASIN Ph. – «L'analyse canonique généralisée et l'analyse des tableaux évolutifs», thèse de 3^{ème} cycle, Strasbourg, 1985.
- CASIN Ph. et TURLLOT J.C. «Une présentation de l'Analyse Canonique Généralisée dans l'espace des individus», *Revue de Statistique Appliquée*, volume XXXIV, n°3, 1986.
- ECONOMIE EUROPEENE N°55, 1993
- ESCOFIER B. et PAGES J. – «Analyses factorielles simples et multiples», Dunod, 1988.
- HORST P. – «Relation among m sets of variables», *Psychometrica* n°26, 1961.
- KETTENRING R.J. – «Canonical analysis of several sets of variables », *Biometrica* n°58, 1971.

- Mc KEON J.J. – «Canonical analysis : Some relations between canonical correlation, factor analysis, discriminant function analysis, and scaling theory», *Psychometric Monograph* N° 13, 1966.
- PERNIN M.O. – «Contribution à la méthodologie d'analyse de données longitudinales», Thèse de Doctorat, Lyon, 1986.
- PONTIER J., DUFOUR A. M. et NORMAND M. «Le modèle euclidien en analyse des données », Editions Ellipses, 1990.
- PONTIER J. et NORMAND M. - «A propos de généralisation de l'analyse canonique», *Revue de Statistique Appliquée*, XL, n° 1, 1992.
- PONTIER J. et PERNIN M.O. «Relations entre analyse canonique complète et méthode; LONGI», *Revue de Statistique Appliquée*, XXXVII, n°4, 1989.
- ROMEDER J.M. - «Méthodes et programmes d'analyse discriminante», Dunod, 1973.
- SAPORTA G. - «Liaisons entre plusieurs ensembles de variables et codage de données qualitatives», Thèse de 3^{ème} cycle, Université de Paris VI, 1975
- SAPORTA G. - «Méthodes exploratoires d'analyse de données temporelles», cahiers du BURO n°37-38, 1981.

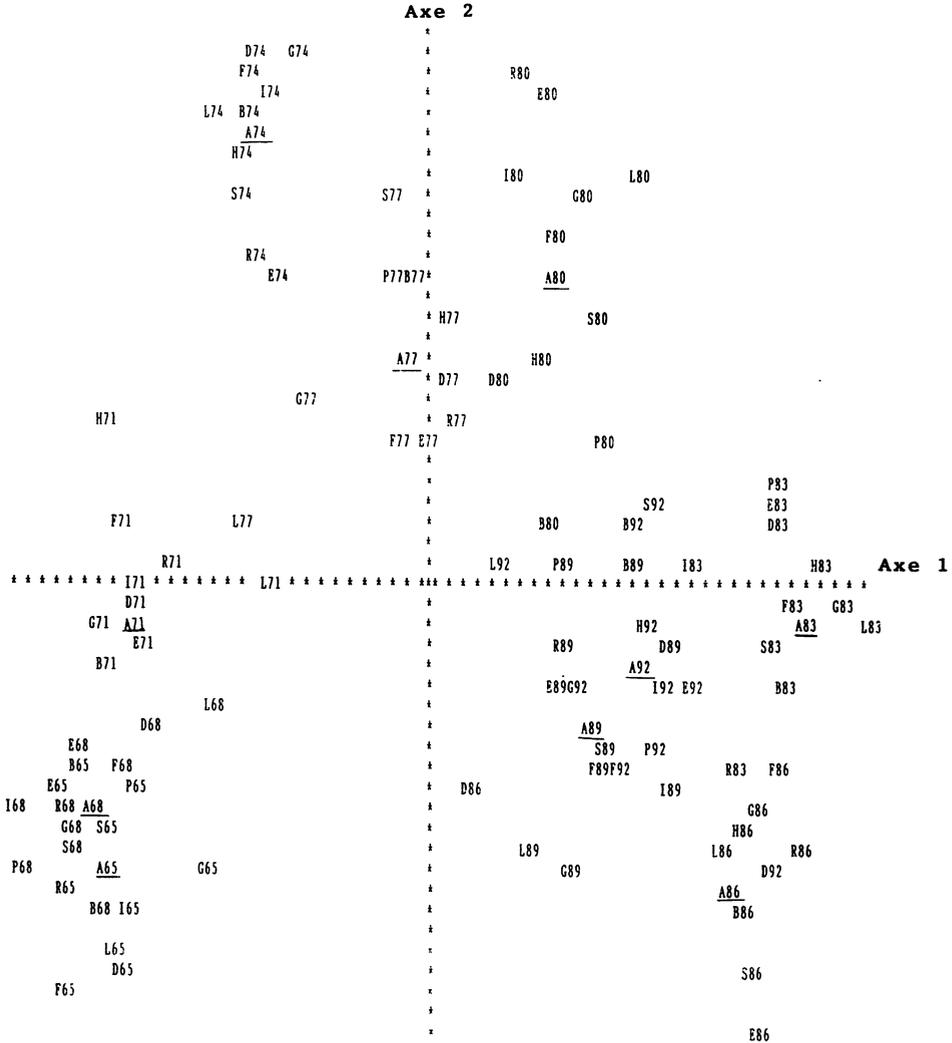
ADTE : ANALYSE DES GROUPES-PAYS



Point vu et, entre parenthèses, point caché

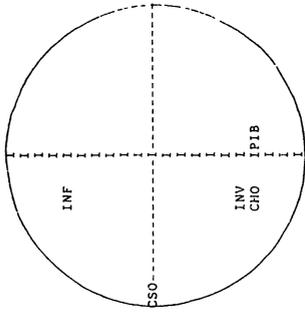
R65 (B80), E77 (S80), F80 (H83), D77 (D86),
 P71 (P86), D77 (D89), I86 (I89), L80 (H89),
 B86 (R89), L89 (B92), H74 (F92), S86 (E92),
 R65 (R92)

ADTE : ANALYSE DES GROUPES-INSTANTS

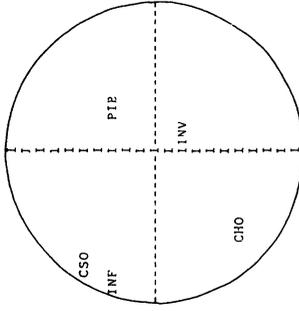


Point vu et, entre parenthèses, point caché

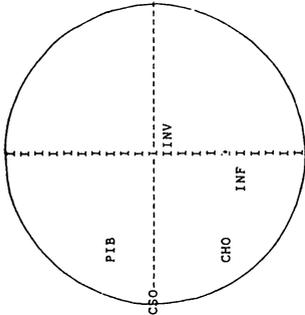
B71 (S71), D80 (I77), B86 (I86), S68 (H65),
 D65 (H68), C89 (H89), E71 (P71), B74 (P74),
 S86 (P86), E92 (R92).



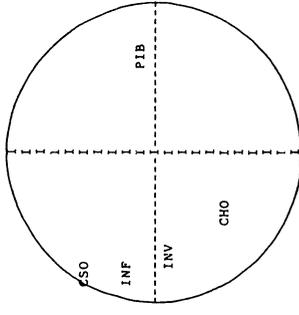
année 74



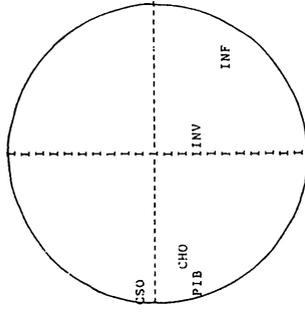
année 86



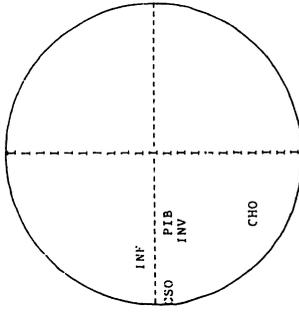
année 71



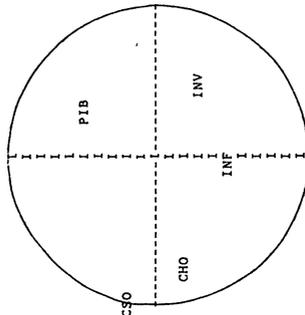
année 83



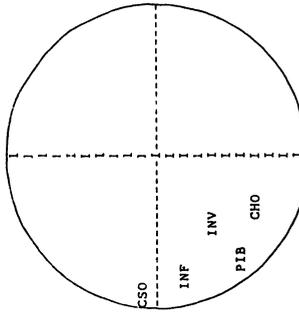
année 68



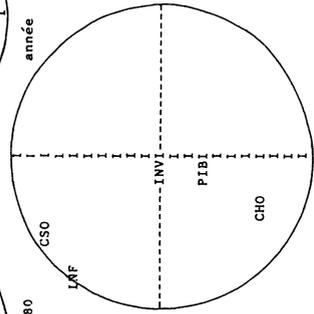
année 80



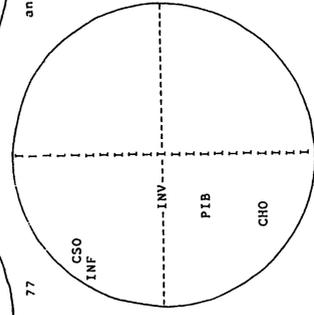
année 65



année 77



année 92



année 89

ADTE: ANALYSE DES GROUPES-PAYS

**ADTE: ANALYSE DES
GROUPES-INSTANTS**

