

# STATISTIQUE ET ANALYSE DES DONNÉES

A. CARLIER

## **Application de l'analyse factorielle des évolutions et de l'analyse intra-périodes**

*Statistique et analyse des données*, tome 10, n° 1 (1985), p. 27-53

[http://www.numdam.org/item?id=SAD\\_1985\\_\\_10\\_1\\_27\\_0](http://www.numdam.org/item?id=SAD_1985__10_1_27_0)

© Association pour la statistique et ses utilisations, 1985, 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/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/>

APPLICATION DE L'ANALYSE FACTORIELLE  
DES EVOLUTIONS ET DE L'ANALYSE INTRA-PERIEDES

A. CARLIER

Laboratoire Modèles et Logiciels en Analyse des Données  
GRECO CNRS 59  
Université Paul Sabatier - 31062 TOULOUSE CEDEX

**Résumé:** *Les méthodes que l'on applique ici sont des analyses en composantes principales locales utilisant une notion de graphe dont les sommets sont les individus. Selon le choix du graphe, on obtient l'analyse factorielle des évolutions ou l'analyse intra-période. Plusieurs choix de codages, métriques et systèmes de poids, adaptés aux tables de contingence à 3 entrées et respectant des propriétés d'équivalences distributionnelles sont utilisés sur le jeu de données.*

**Abstract:** *The methods considered in this paper are local principal components analysis using a graph notion defined on set of individuals. Two different choices of graph lead respectively to factorial analysis of evolutions and intra-period analysis. Different choices of coding, weighting and metrics, which respect the property of "équivalence distributionnelle" are considered on the data set.*

**Mots clés:** *Analyse en Composantes Principales locale, Analyse des évolutions, Intrastructure, Tables de contingence à 3 entrées, Equivalence distributionnelle.*

0 - INTRODUCTION

Les méthodes utilisées ici ont été proposées par L.LEBART pour la

prise en compte de structure de contiguïté, mais dans un cadre assez général, sous le nom d'analyse factorielle locale, et par Y. LE FOLL pour l'étude des données évolutives (AFE). Plus tard Y. LE FOLL en propose une généralisation sous le nom d'analyse factorielle pondérée. Elles remplacent les notions usuelles de covariances et d'inertie par des notions locales permettant d'effectuer des représentations ou des résumés de données, en choisissant de privilégier certains aspects de la structure des données. Ainsi l'analyse factorielle locale permet en introduisant la notion de contiguïté géographique, de mettre en évidence des combinaisons linéaires de variables les plus stables localement (ou au contraire les plus instables). Dans l'AFE, on remplace la notion de contiguïté géographique par une notion de contiguïté temporelle.

I - ASPECTS METHODOLOGIQUES

1 - Les analyses locales ou pondérées

Ici  $L$  indice l'ensemble des individus ( $L = \{1, \dots, n\}$ ),  $J$  l'ensemble des variables ( $J = \{1, \dots, p\}$ ). Le formalisme utilisé est celui de l'analyse en composantes principales (ACP), avec pour l'essentiel les notations de CAILLEZ et PAGES. Le tableau de données  $Y = (y_l^j)_{l \in L, j \in J}$  a pour  $j$ -ème colonne la variable  $Y^j \in \mathbb{R}^{n1}$  et pour  $l$ -ième ligne le transposé du vecteur  $y_l \in \mathbb{R}^p$ .

La matrice  $D = \text{diag}(p_l, l \in L)$  caractérise le système de poids des lignes ( $p_l > 0 \forall l \in L$ ;  $\sum_{l \in L} p_l = 1$ ) et  $\mathcal{M} = \{(y_l, p_l), l \in L\}$  est le nuage pondéré des vecteurs

individus  $y_l$ ; on note  $M$  la matrice définissant le produit scalaire de  $\mathbb{R}^p, \|\bullet\|$

la norme associée,  $I(\mathcal{M})$  l'inertie du nuage  $\mathcal{M}$  par rapport à son barycentre

$Tr$  la forme linéaire trace. Enfin on note  $V$  la matrice des variances-cova-

riances des variables  $y^j (j \in J)$ , dont l'élément  $(j, k)$  est  $\text{cov}(y^j, y^k)$ .

Alors les notions d'inertie et de covariances locales s'obtiennent à partir des expressions (1) et (2) respectivement:

$$I(\mathcal{M}) = \sum_{(l, l') \in L \times L} p_l p_{l'} \|y_l - y_{l'}\|^2 \tag{1}$$

$$\text{cov}(y^j, y^k) = \sum_{(l, l') \in L \times L} p_l p_{l'} (y_l^j - y_{l'}^j) (y_l^k - y_{l'}^k) \tag{2}$$

en ne sommant les mêmes quantités que sur certains couples  $(1,1') \in L \times L$ , l'ensemble des couples  $U$  étant l'ensemble des arêtes d'un graphe non orienté  $G=(L,U)$  d'ensemble de sommets  $L$ . L'analyse factorielle pondérée de  $Y$ . LE FOLL généralise encore l'analyse locale, puisqu'elle permet de remplacer le poids  $p_1 p_{1'}$ , (si  $(1,1') \in U$ ) ou 0 (si  $(1,1') \notin U$ ) du couple  $(1,1')$  par un poids  $p_{11'}$ , quelconque.

Ainsi ces méthodes donnent à l'utilisateur de nouvelles possibilités pour adapter l'outil "analyse factorielle" à la structure des données et aux objectifs de l'analyse. En plus du choix du codage (le tableau  $Y$ ), de la mesure de dissemblance entre lignes (la métrique  $M$  de  $\mathbb{R}^P$ ), de la pondération des individus (définissant une métrique  $D$  de  $\mathbb{R}^n$ ), elles donnent à l'utilisateur le choix du graphe  $G$ , ou plutôt celui de l'ensemble des arêtes  $U$ . Ce choix permet de sélectionner les couples d'individus sur lesquels la variation des variables doit être analysée.

On appelle ici ACP locale de  $(Y,M,D,U)$  l'ACP du tableau  $Y$  dans  $(\mathbb{R}^P, M)$  et  $(\mathbb{R}^n, D)$  pour le graphe  $G=(L,U)$ . On obtient alors les expressions (3) et (4) de l'inertie locale  $I_U(\mathcal{M})$  et de la covariance locale  $cov_U(y^j, y^k)$ :

$$I_U(\mathcal{M}) = \sum_{(1,1') \in U} p_1 p_{1'} \|y_1 - y_{1'}\|^2 \quad (3)$$

$$cov_U(y^j, y^k) = \sum_{(1,1') \in U} p_1 p_{1'} (y_1^j - y_{1'}^j)(y_1^k - y_{1'}^k) \quad (4)$$

Si  $X$  (resp.  $x^j$ ) est le tableau (resp. la variable) centré associé à  $Y$  (resp.  $y^j$ ), ces notions locales s'obtiennent encore, à partir des expressions usuelles:

$$I(\mathcal{M}) = \text{Tr}(VM) \quad (\text{avec } V = {}^t X D X)$$

$$cov(y^j, y^k) = {}^t x^j D x^k$$

en remplaçant la métrique  $D$  de l'espace  $\mathbb{R}^n$  par le semi-métrique  $D_U$  définie par:

$$D_U = \sum_{(1,1') \in U} p_1 p_{1'} \Delta_{11'} \quad (5)$$

où  $\Delta_{11'}$ , carrée d'ordre  $n$ , est la matrice contraste définie par:

$$\begin{aligned}
 (\Delta_{11'})_{r,s} &= 1 && \text{si } r = s = 1 && \text{ou si } r = s = 1' \\
 &= -1 && \text{si } r = 1 \text{ et } s = 1' \text{ ou si } r = 1' \text{ et } s = 1 \\
 &= 0 && \text{sinon.}
 \end{aligned}$$

Pour cette semi-métrie, la norme d'une variable est nulle si et seulement si la variable est constante sur toute partie connexe définie par le graphe. Cette semi-métrie ne nécessitant plus le centrage des variables, les expressions de l'inertie et de la covariance locale deviennent:

$$I_U(M) = \text{Tr}(V_U M) \quad \text{avec } V_U = {}^t Y D_U Y \quad (6)$$

$$\text{cov}_U(y^j, y^k) = {}^t y^j D_U y^k. \quad (7)$$

Pour réaliser l'ACP locale de  $(Y, M, D, U)$ , on peut réaliser l'ACP au sens "classique" de  $(Y, M, D_U)$  (cf. [2] ou [7]).

## 2 - L'analyse factorielle des évolutions

### 2.1 Généralités

Dans la suite, on considère un nouvel ensemble d'individus indicé par  $I (I = \{1, \dots, n\})$  et un ensemble d'instantants  $K (K = \{1, \dots, nk\})$ . Le cube de données  $(x_{ijk})_{(ijk) \in I \times J \times K}$  a pour tranche d'indice  $k$  le tableau  $X_k$  de  $n$  lignes et  $p$  colonnes. On lui substitue le tableau  $Y$  obtenu par empilement vertical de tableaux  $Y_k = T(X_k)$ , de mêmes dimensions que  $X_k$ , et qui peuvent être selon le codage choisi :

- les tableaux  $X_k$  eux-mêmes (codage par effectif)
- les tableaux des profils-lignes de  $X_k$  (codage par profil-ligne)
- d'autres tableaux déduits de  $X_k$ .

L'indice  $l$  d'une ligne de  $Y$  s'écrit  $l = (i, k) \in I \times K$  et on appelle trajectoire de l'individu  $i$  l'ensemble des positions successives de  $y_l$  quand  $l$  décrit, dans l'ordre imposé par le temps, l'ensemble  $\{i\} \times K$ .

Nous étudions dans le paragraphe suivant les choix de graphes  $G$  adaptés

à l'étude de l'évolution des vecteurs individus  $y_{(i,k)}$  au cours du temps.

## 2.2. Le choix du graphe G

Pour décrire les trajectoires des vecteurs associés aux individus-instants  $(i,k)$  au cours du temps, on effectue l'ACP locale de  $(Y,M,D,U)$ , ou  $U$  peut être défini comme:

$$U = \{((i,k), (i,k')), i \in I, k \in K, k' \in K, r_1 \leq k-k' \leq r_2\}$$

avec  $r_1$  et  $r_2$  entiers tels que  $0 < r_1 \leq r_2 < nk$ .

Si  $r_2 < nk-1$ , on ne tient pas compte des variations des variables sur un même individu à des instants extrêmes. Si on choisit  $r_1 > 1$ , on ne tient pas compte des variations entre instants contigus, pour affaiblir l'influence d'un éventuel bruit au profit d'un effet tendanciel plus régulier. Par la suite, nous choisissons  $r_1 = r_2 = 1$ , soit:

$$U = \{((i,k), (i,k+1)), i \in K, k=1, \dots, nk-1\}.$$

En effet, dans notre application, les "instants" successifs sont espacés de 4 années, et les variations, de grande amplitude, sont peu susceptibles d'être affectées par un bruit.

Pour l'étude de l'évolution globale de l'ensemble des lignes  $i$  au cours du temps, nous pouvons représenter, en projection sur les plans factoriels précédemment déterminés, les barycentres  $g_k$  des sous-nuages instantanés

$$M_k = \{(y_i, p_i), i \in I \times \{k\}\}.$$

## 2.3. L'AFE sur table de contingence

### 2.3.1 Généralités

Une telle analyse a été plus spécialement étudiée en ([2]). La table de contingence à 3 entrées a pour élément  $x_{ijk}$  le nombre d'individus appartenant à la catégorie  $i$  de  $I$  et  $j$  de  $J$  à l'instant  $k$  ( $k \in K$ ). L'analyse effectuée ici est dissymétrique: on étudie l'évolution des lignes ou des colonnes du tableau  $X_k = (x_{ijk})_{(i,j) \in I \times J}$  (par la suite on dispose le tableau de sorte que l'analyse effectuée porte sur l'évolution de ses lignes). Cette étude peut ne porter que sur l'évolution structurelle des lignes ou aussi sur leurs évolutions quantitatives. Mais dans tous les cas on souhaite

que l'analyse vérifie une propriété d'équivalence distributionnelle analogue à celle de l'analyse des correspondances.

### 2.3.2. L'équivalence distributionnelle

Sur un tel cube de données, la propriété de double\* équivalence distributionnelle peut s'énoncer ainsi(cf. [2]).

*Si le cube de données  $(x_{ijk})$  possède deux tranches d'indices  $i_1$  et  $i_2$  (resp.  $j_1$  et  $j_2$ ) proportionnelles, où  $i_1$  et  $i_2$  (resp.  $j_1$  et  $j_2$ ) sont deux éléments distincts de  $I$  (resp.  $J$ ), on ne modifie pas les résultats de l'analyse en remplaçant les deux tranches par leur somme.*

Réunir deux tranches  $i_1$  et  $i_2$  se traduit par le remplacement des sommets  $l_1=(i_1,k)$  et  $l_2=(i_2,k)$  par un seul sommet  $l=(i,k), \forall k \in K$ . L'ensemble  $U$  des arêtes est lui aussi modifié. En AFE, cet ensemble définit sur  $L$  les parties connexes  $\{i\} \times K$  associées à chaque tranche, et les expressions (3) et (4) se décomposent en somme de contributions de chaque tranche  $i$  ( $i \in I$ ). La conservation de l'inertie ou de la covariance par réunion de tranches est obtenue lorsque la proportionalité des tranches implique celle de leurs contributions à l'une ou l'autre de ces expressions (voir § 2.3.3). C'est donc le choix du codage et du système de poids qui intervient ici. L'équivalence distributionnelle sur les tranches  $j$  de  $J$  résulte, comme en analyse factorielle, du choix de la métrique. Cette métrique est choisie pour assurer à la contribution d'une colonne  $j$  à la distance entre deux points d'être homogène à l'effectif de la colonne.

\* La triple équivalence distributionnelle n'a pas de sens pour le type de données traitées ici. On peut imaginer des cas (par exemple: nombre de naissances par canton et par catégorie socio-professionnelle sur des intervalles de temps différents) où elle pourrait en avoir un. Mais elle n'est pas considérée ici.

### 2.3.3. Choix du codage et du système de poids

Choix 1: codage en profil ligne

On note  $x_{(i,k)}$  le transposé de la  $i$ -ème ligne du tableau  $X_k$ .

Si on s'intéresse à l'évolution de la structure de ces lignes, quand  $k$  décrit  $K$ , on est amené à utiliser le codage en profil ligne. Le tableau  $Y$  traité, noté  ${}_1Y$ , admet pour ligne  $(i,k)$ :

$${}^t y_{(i,k)} = {}^t x_{(i,k)} / x_{i.k} \quad \text{avec } x_{i.k} = \sum_{j \in J} x_{ijk} .$$

Si on écrit l'inertie totale comme somme des contributions  $ct_j^I$  des tranches

$i$  de  $I$ , on a:

$$I_U(m) = \sum_{i \in I} ct_j^I \quad \text{avec :}$$

$$ct_j^I = \sum_{\substack{(k,k') \in K^2 \\ ((i,k),(i,k')) \in U}} p_{ik} p_{ik'} \| y_{(i,k)} - y_{(i,k')} \|^2$$

Pour obtenir l'homogénéité de cette contribution à l'effectif de la tranche  $i$ , on aurait pu prendre comme  $Y$ . LE FOLL(1972):

$$p_{i,k} = \sqrt{x_{i..} / x_{...}} \quad \text{avec } x_{i..} = \sum_{j,k} x_{ijk} \quad \text{et } x_{...} = \sum_{i,j,k} x_{ijk} .$$

Dans la suite nous utilisons un système de poids un peu différent:

$$p_{i,k} = \sqrt{x_{i.k} / x_{...}} \quad \text{avec } x_{i.k} = \sum_j x_{ijk} .$$

Avec un tel choix, le poids affecté au couple  $((i,k)(i,k'))$  dépend des instants  $k$  et  $k'$  considérés.

Choix 2: codage en effectif

Ici on cherche à étudier les évolutions structurelles et quantitatives des lignes des tableaux  $X_k$ . Le tableau  ${}_2Y$  s'obtient par simple



empilement des tableaux  $X_k$ . Les lignes  $x_{(i,k)}$  et  $x_{(i',k)}$  associées à deux tranches proportionnelles ne sont plus confondues, et leurs trajectoires sont affines (et même homothétiques si aucun centrage n'est effectué). Ici la quantité  $\|y_{(i,k)} - y_{(i',k)}\|^2$  est homogène au carré de l'effectif de la tranche  $i$ . Pour satisfaire la propriété d'équivalence distributionnelle, on peut prendre comme système de poids:

$$p_{i,k} = \sqrt{x_{i..} / x_{i..}} \quad (i,k) \in I \times K$$

ou encore:

$$2 p_{i,k} = \sqrt{x_{i..} / x_{i..k}} \quad (i,k) \in I \times K$$

On choisit le deuxième système, et on note  ${}_2D$  la matrice diagonale associée.

Choix 3: codage en profil tranche:

On appelle profil de la tranche  $i$  du cube  $(x_{ijk})$  le tableau à 2 dimensions  ${}_3Y_k = (x_{ijk} / x_{i..})_{(j,k) \in J \times K}$ , où l'indice  $k$  est l'indice de ligne.

Alors comme pour le choix 2, l'AFE effectuée sur le tableau  ${}_3Y$ , obtenu en empilant verticalement les sous-tableaux  ${}_3Y_k$ , permet d'étudier les évolutions structurelles et quantitatives des lignes du tableau  $X_k$ , quand  $k$  décrit  $K$ . Mais ici la taille moyenne du canton n'intervient pas dans le codage.

La trajectoire  $\{{}_3Y_{(i,k)}, (i,k) \in I \times K\}$  se déduit de la trajectoire obtenue dans le codage en effectif par une homothétie de centre 0 et de rapport  $\frac{1}{x_{i..}}$ .

Soit encore:

$${}_3Y = D_I^{-1} {}_2Y \quad \text{avec } D_I = \text{diag}(x_{i..}, (i,k) \in I \times K).$$

Alors en choisissant les poids  ${}_3p_{i,k}$  tels que :

$${}_3p_{i,k} = x_{i..} \quad {}_2p_{i,k} = x_{i..} \sqrt{x_{i..} / x_{i..k}}$$

Les matrices de variances-covariances locales sont identiques pour les choix

2 et 3. En choisissant une même métrique dans  $\mathbb{R}^P$ , on obtient pour les deux analyses:

- les mêmes valeurs propres et vecteurs propres
- les mêmes représentations dans l'espace des variables
- les représentations des trajectoires qui se déduisent par cette homotéthis.

#### 2.3.4. Le choix de la métrique de $\mathbb{R}^P$

Nous ne revenons que brièvement sur les choix effectués ici, qui assurent la propriété d'équivalence distributionnelle sur les tranches  $j$  de  $J$ .

Les métriques  $M$  choisies étant à priori diagonales, la contribution d'une variable à la variance totale  $\text{Tr}(V_U M)$  s'écrit  $m_{jj} \text{var}_U(Y^j)$ .

Pour donner à chaque variable une importance proportionnelle à sa dispersion locale, et homogène à l'effectif de la colonne, nous choisissons la métrique:

$$M = D_{1/\sigma} = \text{diag} \left( \frac{1}{\sigma_{jU}}, j \in J \right) \quad \text{avec} \quad \text{var}_U(Y^j) = (\sigma_{jU})^2.$$

Ainsi, la contribution d'une variable à l'analyse est égale à son écart-type local  $\sigma_{jU}$ .

### 3- L'analyse intra-période(AIE)

#### 3.1 Généralités

Cette analyse, proposée par Y. LE FOLL en 1973 sous le nom d'analyse en composantes principales essentielle, est encore une ACP locale. Mais ici le graphe  $U$  est choisi pour mettre en évidence la structure interne des sous-nuages  $\mathcal{M}_k$ . On prendra ici:

$$U_I = \left\{ ((i,k)(i',k)), \forall k \in K, \forall (i,i') \in I \times I, i < i' \right\}$$

(un vecteur individu à l'instant  $k$  est relié à tous les vecteurs individus au même instant). L'inertie locale du nuage  $\mathcal{M}$  est alors la somme des inerties totales de chaque sous-nuage. On retrouve encore une des deux ACP de la double ACP de Bouroche. On pourra alors représenter chaque sous-nuage  $\mathcal{M}_k$  sur le plan de compromis, et représenter le nuage  $\bar{\mathcal{M}}$  des barycentres des trajectoires de chaque canton.

### 3.2. Choix de codages, métriques et poids.

Pour obtenir la propriété de double équivalence distributionnelle, il est ici nécessaire de choisir un codage invariant quand une tranche  $i$  est remplacée par une tranche proportionnelle (codage en profil ligne ou en profil tranche). Nous avons ici utilisé le codage en profil ligne (étude de la structure des cantons indépendamment de leur taille), avec des poids proportionnels aux effectifs des lignes. La métrique utilisée ici est la métrique  $D_{1/\sigma}$  définie en 2-3-4, et calculée à partir des variances locales des variables.

## 4. Quelques éléments d'interprétation d'une ACP locale

### 4.1. Les contributions

La métrique  $M$  étant diagonale, on étudie dans chaque cas les contributions relatives des variables  $j$  à l'inertie totale (notées  $ct_j^J$ ) et à la variance locale du facteur (ou composante principale)  $l$ , notée  $ct_j^l$ . En notant  $s_1^j$  la saturation du facteur  $l$  dans la variable  $y^j$  et  $\lambda_l$  la variance locale du facteur  $l$ , on a :

$$ct_j^J = m_{jj} \text{ var}_U(y^j) / \text{Tr}(V_U M)$$

$$ct_j^l = m_{jj} (s_1^j)^2 / \lambda_l$$

L'expression (3) de l'inertie montre que la décomposition de l'inertie se fait par arêtes, et non par élément  $(i,k)$  de  $I \times K$ .

Il peut être utile de calculer les contributions des sous-ensembles connexes

déterminés par le graphe, soit les contributions des tranches  $i$  en AFE, ou des instants  $k$  en AIE. Ces dernières contributions, non calculées par la version actuelle du logiciel<sup>(\*)</sup> utilisé, ne sont pas interprétées ici.

#### 4.2. Les corrélations locales et totales.

Les composantes principales des ACP locales sont non corrélées localement, mais en général, corrélées au sens de la métrique diagonale des poids. On peut alors étudier:

-sur les cercles de corrélations locales, les corrélations locales variables facteurs. En AFE, on étudie ainsi quelles sont les variables qui évoluent ensemble ou au contraire qui évoluent en sens contraire. En AIE, on étudie comment en moyenne, les variables structurent l'ensemble des individus à chaque instant.

-sur les cercles des corrélations totales, les corrélations variables facteurs au sens classique. Cela permet d'interpréter les composantes principales des individus à chaque instant, sans tenir compte de la structure définie par le graphe. En AFE, on peut ainsi interpréter les positions relatives des trajectoires entre elles, en AIE les positions relatives des nuages  $\mathcal{M}_k$  entre eux. Remarquons que les facteurs étant en général corrélés au sens classique, ils ne sont pas orthogonaux. La lecture des corrélations variables facteurs se fait ici par projection orthogonale(cf.figure 1) -

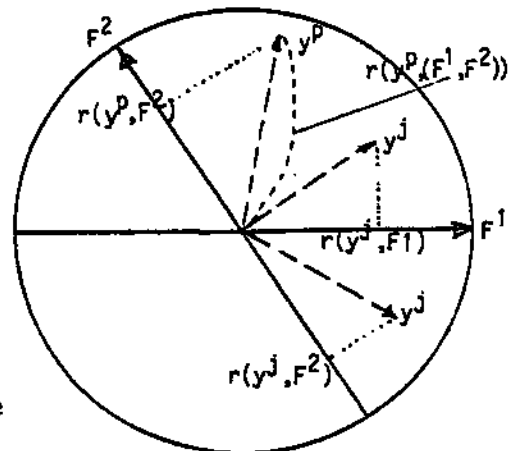


Figure 1

Cercle des corrélations totales

(\*) Logiciel "Enseignement et pratique de l'Analyse des Données" implanté au Centre Interuniversitaire de Calcul de Toulouse.

4.3. Un indice de variabilité

On note  $d_j$  la fréquence de la tranche  $j$  du cube  $(x_{ijk})$ , soit:

$$d_j = x_{.j} / x_{...} \quad \text{avec } x_{.j} = \sum_{i,k} x_{ijk}$$

On utilise dans la suite l'indice de variabilité proposé en ([2]) :

$$r_j = \frac{ct_j^j}{d_j} \quad , j \in J, \quad \text{dont la moyenne pondérée est } 1 .$$

Il mesure la variabilité de la variable à travers les choix (codages, métrique, système de poids) effectués par l'analyse. Pour la métrique du  $\chi^2$ , cet indice vaut:

$$r_j = (\text{var}_U(y^j) / (d_j)^2) / \text{Tr}(V_{JM})$$

Pour la métrique  $D_{1/\sigma}$ , on obtient des coefficients proportionnels à la racine des précédents.

II- APPLICATION A L'ETUDE DE L'EVOLUTION DE LA POPULATION DU LANGUEDOC-ROUSSILLON:

1 - Analyse factorielle des évolutions

1.1 - Etude de l'évolution structurelle de la population

1.1.1 - Rappel des choix effectués

On étudie ici l'évolution structurelle de la population des cantons, indépendamment des variations de leurs effectifs. Rappelons les choix de  $(Y, M, D, U)$  effectués ici:

$Y = {}_1Y$  a pour ligne  $(i,k)$  :  $t_{x(i,k)/x_{i,k}}$

$M = {}_1M$  a pour élément diagonal  $j$  :  $1 / \sigma_{jU}$

$D = {}_1D$  a pour élément diagonal  $(i,k)$  :  $p_{i,k} = \sqrt{\frac{x_{i,k}}{x_{...}}}$

et  $U = \{(i,k), (i,k+1)\}, i \in I, k=1, \dots, nk-1\}$

### 1.1.2. - Indice de variabilité et contribution

Le tableau 1 donne en pourcentage, par CSP  $j$ , les contributions  $ct_j^J$  à l'inertie totale, la fréquence  $d_j$  et le coefficient de variabilité  $r_j$ .

CSP	EA	OA	AC	PL	CM	EM	OU	SE	CP
$ct_j^J$	20	21	4	7	10	11	20	4	3
$d_j$	20	15	12	3	7	8	30	3	2
$r_j$	1	1,4	0,3	2,3	1,4	1,4	0,7	1,3	1,5

Tableau 1

Les variables les plus stables sont celles dont la proportion dans la population reste constante. Ce sont les catégories Ouvriers (0,7), mais surtout "Artisans Commerçants" (0,3). Les variables les moins stables sont celles qui ont le plus évolué en pourcentage: on trouve les ouvriers agricoles, les cadres moyens et les employés (1,4), les autres catégories (1,5) et surtout les professions libérales (2,3). Les contributions des CSP dépendent à la fois de leur variabilité, et de leur effectif. Ainsi les catégories agricoles (35% de la population) représentent 41% de la variation totale, les catégories tertiaires (PL, CM, EM, SE, 21% de la population) en représentant 38%. La catégorie Ouvriers (30% de la population) n'apporte que 20% de cette variation. On trouve ensuite les Artisans Commerçants (12% de la population, 4% de la variation) et les autres catégories (respectivement 2% et 3%). Le tableau 2 donne

les pourcentages de chaque facteur à la variation (ou à l'inertie) totale, et celles des CSP à la variance locale des facteurs.

$1 = \frac{1}{2} \text{CSP}$	EA	OA	AC	PL	CM	EM	OU	SE	CP	
1	62	27	30	0	6	11	13	11	2	0
2	20	0	1	1	10	11	10	66	1	0
3	5	50	40	0	1	1	4	0	0	4
4	4	0	2	84	1	0	8	0	4	1
5	3	0	1	4	0	0	0	0	24	70

Tableau 2

### 1.1.3. Interprétation des corrélations locales

Le cercle des corrélations locales dans le plan des facteurs 1 et 2 (figure 2) donne les grands traits de l'évolution de la structure des cantons (82% de la variation totale). Il montre que les catégories agricoles varient ensemble, ainsi que les catégories tertiaires PL, CM, EM. Une baisse du secteur primaire s'accompagne d'une hausse des secteurs secondaires et/ou tertiaires, mais les évolutions des CSP tertiaires et secondaires ne sont pas liées. Enfin les évolutions des catégories CP et AC n'apparaissent pas liées aux précédentes.

L'axe 1 (62% de l'inertie) oppose les catégories agricoles (corrélation locale inférieure à  $-0,93$ ) aux catégories tertiaires PL, CM et EM (corrélation supérieure à  $0,73$ ) et ouvrières ( $r_U(OU, F^1) = 0,58$ ). Ces 6 catégories contribuent à 98% de sa variance locale. L'axe 2 (20% de l'inertie) est l'axe du secteur secondaire ( $r_U(F^2, OU) = -0,81$ , et OU contribue à 66% de la variance de  $F^2$ , dont il marque en partie l'opposition avec le tertiaire (corrélations locales avec PL, CM et EM supérieure à  $0,43$ ). Les axes 3, 4 et 5 représentent respectivement 5, 4 et 3% de l'inertie. L'axe 3 différencie les évolutions des catégories exploitants et ouvriers agricoles (qui contribuent à 90% de sa variance) l'axe 4 traduit celle des artisans commerçants (84% de  $var_U F^4$ ,  $r_U(F^3, AC) = 0,95$ ) et l'axe 5 celle des autres catégories (70% de  $var_U F^5$ ,  $r_U(F^5, AC) = 0,85$ ) qu'il oppose en partie à celle du personnel de service ( $r_U(F^5, SE) = -0,46$ ).

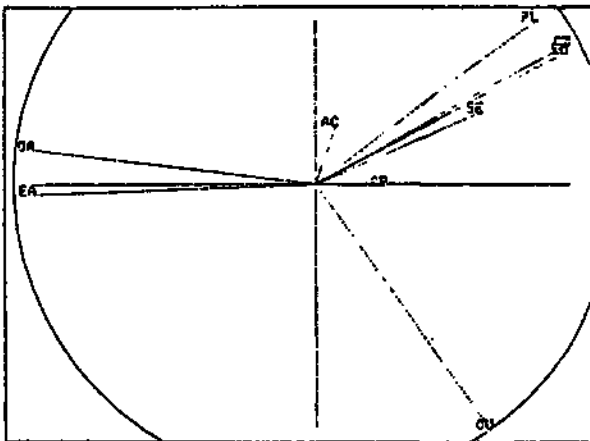


Figure 2

Cercle des corrélations locales 1,2

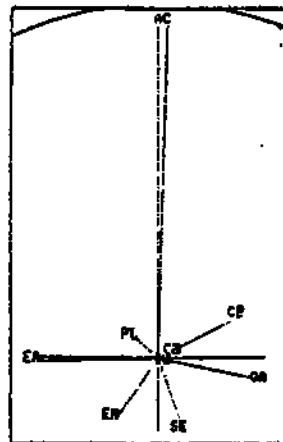


Figure 3

Cercle des corrélations locales 3,4

1.1.4. Interprétation des corrélations totales

L'étude des corrélations totales (cf. figure 5) permet d'identifier les positions relatives de deux vecteurs canton-instant quelconques, et en particulier les positions relatives des trajectoires. L'axe 1 reste un axe d'opposition rural-urbain: l'abscisse de l'origine d'une trajectoire sur l'axe 1 traduit son caractère plus ou moins rural ou urbain. L'axe 2 devient un axe opposant secteur primaire ( $r(F^2, OA) = 0,57$  et  $r(F^2, EA) = 0,47$ ) et secteur secondaire ( $r(F^2, OU) = -0,93$ ).

L'ordonnée des positions d'origine des trajectoires traduit le caractère plus ou moins rural, ou plus ou moins industrialisé des cantons. L'axe 3 est l'axe repérant la composition rurale des cantons ( $r(F^3, EA) = -0,80$ ,  $r(F^3, OA) = 0,31$ ), l'axe 4 reste l'axe des artisans commerçants ( $r(F^4, AC) = 0,98$ ). Enfin l'axe 5 oppose en partie EA et CP à AC et SE (corrélations respectives 0,58; 0,53; -0,54; -0,52).

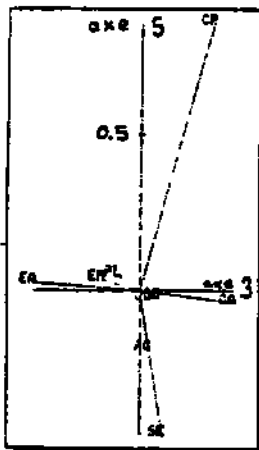


Figure 4

Cercle des corrélations locales 3-5

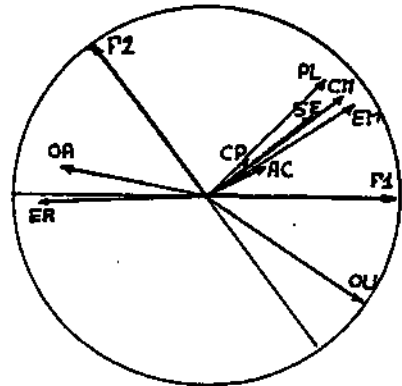


Figure 5

Cercle des corrélations totales 1-2

1.1.5. Etude des trajectoires

Nous donnons (figure 6) l'aspect général de ces trajectoires sur le plan 1-2 et (figure 7) une sélection de certaines d'entre elles, accompagnée (tableau 3) des pourcentages respectifs des catégories primaires secondaires et tertiaires.



Ces trajectoires ont grossièrement la même allure concave. Quelle que soit leur structure d'origine, les cantons subissent l'exode rurale; dans la première période, ils augmentent surtout en pourcentage leur population ouvrière, durant la dernière surtout leur population tertiaire. Pour la



Figure 6

Trajectoires plan 1-2

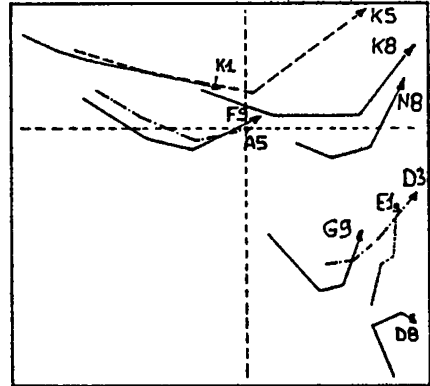


Figure 7

Trajectoire de 10 cantons plan 1-2

période intermédiaire, leur comportement se situe entre les deux.

Ainsi  $K_1$  (canton de l'Hérault) est très rural à l'origine; son évolution est modérée et peu marquée par l'industrialisation.

C.S.P.	EA-OA				CM-EM-PL				OU			
	t=1	2	3	4	1	2	3	4	1	2	3	4
canton												
A5	67	59	47	36	5	7	11	17	15	21	28	30
D3	21	16	11	5	12	14	19	30	55	56	54	50
D8	8	12	7	4	8	11	16	16	76	67	67	69
E1	10	8	6	6	11	15	18	24	64	58	58	50
F9	71	57	49	35	4	7	10	20	14	24	29	28
G9	33	23	18	15	11	9	13	21	47	58	59	52
K1	84	77	71	44	5	7	10	21	3	7	10	23
K5	72	56	37	16	7	12	22	45	9	16	25	21
K8	46	30	16	7	13	19	30	45	22	30	35	29
N8	23	16	9	5	13	15	22	31	34	38	39	31

Tableau 3

Le canton K5 (autre canton de l'Hérault) a une origine aussi fortement rurale, mais son évolution structurelle est beaucoup plus forte et très marquée par le secteur tertiaire. K8 (canton de Montpellier) a une origine très urbaine, il s'industrialise un peu plus et termine avec une forte structure tertiaire. Les cantons A5 (Castelnaudary) et F9 (Uzès), assez ruraux d'origine, ont une évolution plus industrielle. Enfin, les cantons du Gard G9, E1, D3 et D8 ont à l'origine une structure très ouvrière. Les trois premiers ont un pourcentage de population rurale qui continue à décroître ainsi que, pour les périodes 2 et 3, celui de leur population ouvrière. Le dernier est le seul canton qui "gagne" (en fréquence seulement) de la population rurale en première période, d'où son allure un peu particulière.

L'étude des projections sur l'axe 3 (cf. figure 8 : représentation des trajectoires - plan 1-3) permet de différencier les évolutions des CSP rurales exploitants et ouvriers agricoles.

La différenciation de ces deux CSP, dont on a vu qu'elles varient assez bien ensemble ( $r_{U(OA,EA)}=0,77$ ) est en fait assez triviale et provient de l'effet de la structure rurale initiale (agriculture familiale ou industrielle) sur l'exode rural: les cantons qui comportaient un fort pourcentage d'exploitants agricoles ont surtout perdu des exploitants agricoles. Or l'axe 3 est au sens des corrélations totales l'axe qui oppose les exploitants agricoles ( $r(F^3,EA)=-0,8$ ) aux

autres CSP, dont la CSP OA ( $r(F^3,OA)=0,3$ ). Et un canton dont la position initiale sur l'axe 3 est négative a une agriculture familiale dominante. L'exode rural se traduit principalement pour lui par une perte d'exploitants agricoles, d'où l'évolution en faisceau. On peut cependant mettre en évidence les cantons dont le comportement est particulier (H9 de l'Hérault passe d'une proportion relative de 49% d'exploitants agricoles et 51% d'ouvriers agricoles en 1954 à une proportion 62% EA, 38% OA en 1975; K1, du Gard passe de 57% EA - 43% OA à une proportion respective de 71%-29% dans la même période).

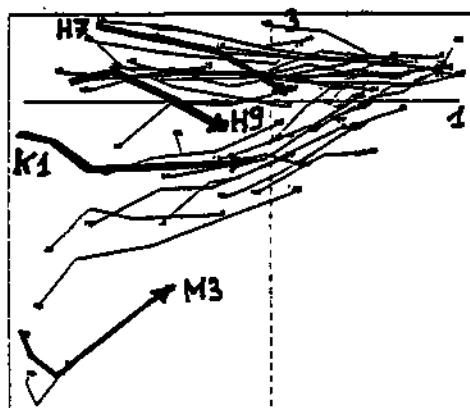


Figure 8

Trajectoires - plan 1-3

La représentation des trajectoires sur le plan 4-5 permet d'étudier les évolutions particulières des catégories AC, CP et, en partie seulement, SE. Nous ne l'avons pas fait ici.

## 1.2 - Interprétation de l'évolution structurelle et quantitative

### 1.2.1 - Considérations générales

Les choix 2 et 3 nous donnent des résultats identiques pour toutes les sorties utilisées ici, à l'exception des représentations des trajectoires, et des corrélations totales, qui permettent d'interpréter leurs positions relatives. C'est pourquoi nous les étudions simultanément. Les contributions  $ct_j^U$  des CSP, leur fréquence  $d_j$ , et leur coefficient de variabilité  $r_j$  figurent dans le tableau 4. Les catégories ont des contributions peu différentes de celles de l'étude précédente, excepté la CSP OU (qui passe de 20 à 24%), AC (de 4 à 6%) et surtout OA (de 21 à 14%).

CSP	EA	OA	AC	PL	CM	EM	OU	SE	CP
$ct_j^U$	18	14	6	7	11	13	24	4	3
$d_j$	20	15	12	3	7	8	30	3	2
$r_j$	0,9	0,9	0,5	2,3	1,6	1,6	0,8	1,3	1,5

Tableau 4

### 1.2.2 - Interprétation des facteurs

Le tableau 5 donne par facteur, les pourcentages d'inertie expliquée, et les contributions des catégories aux facteurs. Le cercle des corrélations locales 1-2 (fig. 9) montre que le premier axe (57% de l'inertie) reste un axe d'opposition rural-urbain. L'axe 2 est un facteur taille, traduisant surtout l'évolution des catégories agricoles (qui contribuent à 74% de sa variance),

k	CSP	EA	OA	AC	PL	CM	EM	OU	SE	CP
1	57	10	13	4	8	16	18	26	4	1
2	19	56	18	12	3	3	3	4	1	0
3	13	3	3	0	11	11	11	59	2	0
4	5	29	59	2	0	0	2	7	0	1
5	2	0	0	3	2	0	2	0	22	71

Tableau 5

et artisans-commerçants (12%). Les axes 3, 4 et 5 s'interprètent comme les axes 2, 3 et 5 de l'analyse précédente (cf. figure 10 et 11): l'axe 3 oppose le secteur secondaire au secteur tertiaire (PL, EM, CM et OU contribuent à 82% de  $\text{var}_U F^3$ ), l'axe 4 différencie les catégories exploitants et ouvriers agricoles (88% de  $\text{var}_U F^4$ ) et l'axe 5 traduit l'évolution de la catégorie CP qu'il oppose en partie à celle de SE (93% au total de  $\text{var}_U F^5$ ).

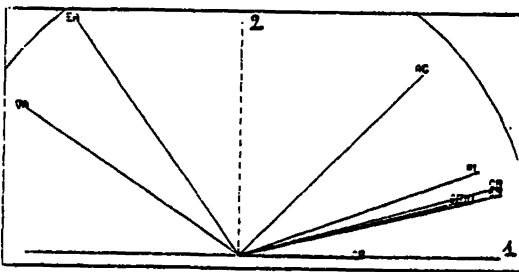


Figure 9

Cercle des corrélations locales 1-2

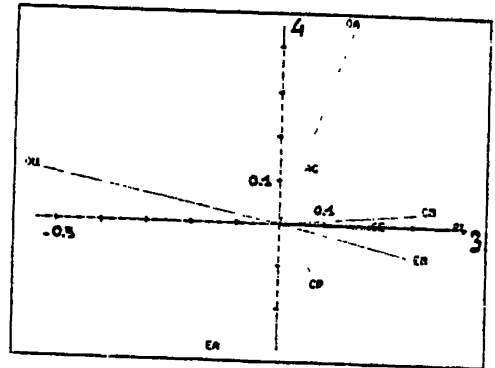


Figure 10

Cercle des corrélations locales 3-4

On avait remarqué dans l'analyse précédente la faible variabilité de AC qui signifiait que l'évolution de cette catégorie était proche de celle de l'ensemble du canton. Avec le codage en effectif un facteur taille s'est inséré en deuxième position, et a pris en compte une bonne part de la variabilité de AC. On retrouve ensuite dans l'ordre les facteurs de l'analyse précédente, à l'exception du quatrième qui caractérisait AC.

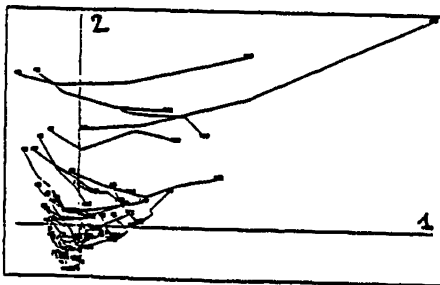


Figure 11.

Trajectoires-choix 2-Plan 1-2

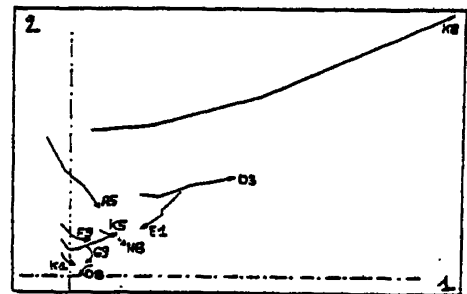


Figure 12

10 trajectoires-choix 2-plan 1-2

### 1.2.3. Interprétation de trajectoires

#### Codage en effectif

Pour interpréter les situations relatives des trajectoires, nous utilisons les corrélations totales (figure 13). Le premier axe est un axe de taille "urbaine", le second un axe de taille générale.

On passe d'une trajectoire de la figure 12 à une trajectoire de la figure 15, par une homothétie dont le centre est l'origine des axes (figure 12) et le rapport est  $\frac{1}{x_{1..}}$  (des centrages distincts, figures 11, 14, sans influence sur l'analyse ont opéré une translation des repères).

Ici on étudie les évolutions simultanément sur les graphes obtenus par les choix 2 et 3: le choix 2 permet d'apprécier les variations structurelles et quantitatives en unité d'effectif, le choix 3 relativement à l'effectif du canton. Les figures 13 et 16 nous donnent les représentations des corrélations totales pour chacun de ces choix. Ainsi pour le choix 2, l'axe 1 oppose les cantons-instants selon l'importance de leur effectif urbain et pour le choix 3 selon l'importance relative de cet effectif.

De même pour le choix 2, l'axe 2 est un axe de taille, alors que pour le choix 3, il oppose les cantons ruraux aux cantons industrialisés.

Sur les figures 14 et 15, les trajectoires dont l'origine est située en haut sont très rurales, peu urbaines et peu industrialisées. Elles subissent un fort exode rural, en général sans un développement similaire des catégories urbaines (K1, F9, A5).

Certains cantons, plus ou moins ruraux à l'origine ont un fort ou un très fort développement urbain. Ainsi K5 subit une très forte variation structurelle et quantitative (de 72 à 16% d'agriculteurs, de 17 à 59% de population tertiaire, de 9 à 21% d'ouvriers; sa population est multipliée par 2). Le canton K8 (Montpellier) d'origine plus urbaine, subit une moindre variation structurelle, mais une plus forte variation quantitative (agriculteurs de 46 à 7%,

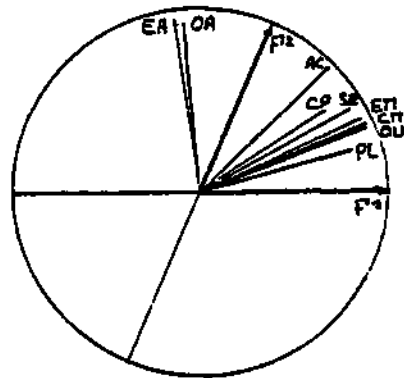


Figure 13  
Cercle des corrélations  
totales-choix 2-plan 1-2

tertiaire de 30 à 62%, ouvriers de 22 à 29%, population multipliée par 2,75).

Dans certains cantons, l'exode rural ne s'accompagne d'aucun développement des catégories urbaines, ou même d'une régression de ces catégories. Ce sont en général des cantons assez urbains (partie droite, figure 14 ou 15) et souvent industrialisés (axe 2 négatif). On trouve parmi ces cantons les cantons du Gard G9, D8 et E1; leurs trajectoires sont soit quasi verticales (G9) ou de pentes négatives (forte régression: D8, E1). Leurs populations urbaines sont multipliées entre 54 et 75 par 0,8 (G9), 0,6 (E1) et 0,5 (D8). Par contre l'autre canton du Gard D3 connaît un fort développement de ces couches urbaines ( $\times 1,8$ ).

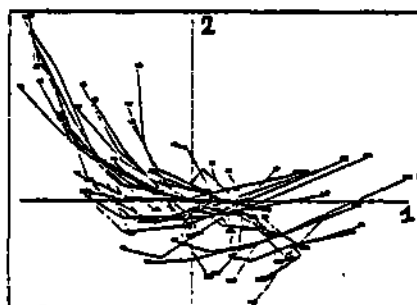


Figure 14

Trajectoires- choix 3 - plan 1-2

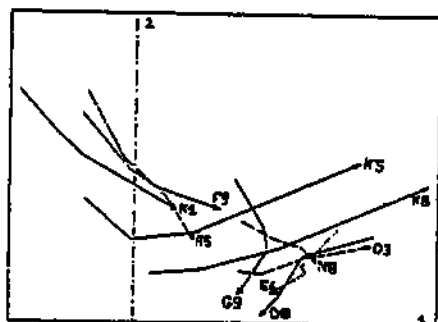


Figure 15

10 trajectoires- choix 3 - plan 1-2

Le plan 3-4 (cf. figures 17 et 18) permet de différencier les évolutions secondaires ou tertiaires des couches urbaines (axe 3) et différencie les évolutions des catégories agricoles (axe 4). Pour le choix 2, l'examen des corrélations totales montre que l'axe 3 oppose aussi les cantons industriels aux cantons tertiaires. On constate pour la plupart des cantons une évolution vers la gauche pendant la première période, une évolution vers la droite de plus forte amplitude pendant la dernière, l'évolution pendant la période intermédiaire étant plus différenciée (par exemple F9 et G9). Une évolution de gauche vers la droite s'interprète comme un accroissement relatif du secteur tertiaire par rapport au secondaire, l'accroissement général de ces catégories étant déjà traduit sur l'axe 2.

Enfin on constate une tendance générale des trajectoires à converger vers

l'axe 4. On retrouve ici le phénomène de différenciation de l'exode rural analysé par le facteur 3 de l'analyse précédente.

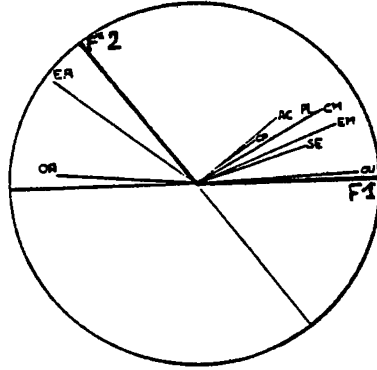


Figure 16  
Cercle des corrélations totales  
choix 3 - plan 1-2

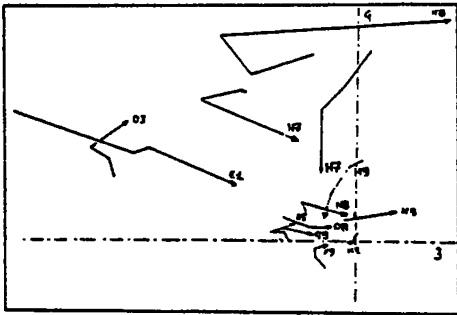


Figure 17  
Trajectoires - choix 2 - Plan 3-4

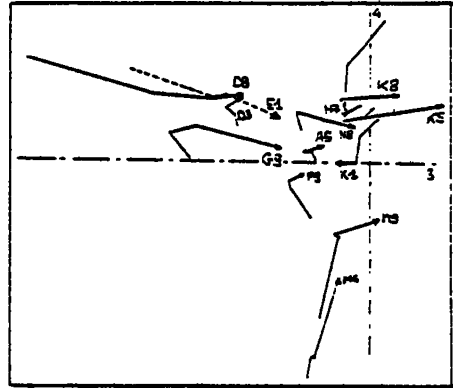


Figure 18  
Trajectoires - choix 3 - Plan 3-4

## 2 - L'ANALYSE INTRA-PERIODE

### 2.1 - Généralités

Les choix utilisés ici sont les mêmes que ceux du paragraphe 1.1, à l'exclusion du choix du graphe:

codage en profil ligne

ponds proportionnels aux effectifs des lignes

métrique  $D_{1/\sigma} = \text{diag} \left( \frac{1}{\sigma_{jU}}, j \in J \right)$

le graphe choisi réunit tous les cantons-instants associés à un même instant.

L'analyse permet d'étudier les principales caractéristiques de différenciation de la structure des cantons. Les plans obtenus sont des plans de compromis qui traduisent au mieux la structure interne de chaque sous-nuage instantané  $M_k = \{y_{(i,k)}, i \in I\}$ .

Le choix de la métrique assure aux catégories socio-professionnelles une contribution proportionnelle à leur écart-type locale (cet écart-type intra est, à un facteur multiplicatif près, la moyenne des écarts-types des variables à l'intérieur de chaque sous-nuage).

## 2.2. Indices de variabilités et contributions

CSP	EA	OA	AC	PL	CM	EM	OU	SE	CP
$ct_j^J$	24	21	7	4	4	6	28	4	2
$d_j$	20	15	12	3	7	8	30	3	2
$r_j$	1,2	1,4	0,6	1,3	0,6	0,7	0,9	1,3	1

Tableau 6

On voit sur ce tableau que les catégories qui varient le plus à l'intérieur des sous-nuages sont les catégories Ouvriers Agricoles-Personnels de service-Professions libérales et Exploitants Agricoles. Celles au contraire qui sont les plus stables sont les Artisans Commerçants, les Cadres Moyens et les Employés. Globalement, les catégories agricoles contribuent à 45% de la variabilité intra des cantons, le secteur tertiaire (PL, CM, EM, SE) à 18% les ouvriers à 28%. Si le secteur tertiaire était sur un bon indicateur pour caractériser l'évolution des cantons, il intervient peu dans leur différenciation entre eux (la contribution passe de 35 ou 38% à 18%).

## 2.3- Interprétation des facteurs

Le tableau des contributions des catégories aux facteurs (tableau 7) et les cercles de corrélations locales (figure 19 et 21) permettent d'interpréter les différents éléments de différenciation des cantons. On donne



aussi la figure 20 où sont représentés les cantons-instants sur le plan de compromis 1-2. On y a ajouté les trajectoires (même si ce n'est pas l'objet de l'étude) et la représentation des cantons moyens ( barycentre de la trajectoire). On remar-

que tout d'abord la mauvaise représentation des variables tertiaires. Leurs faibles variations intra s'accompagnent de faibles corrélations intra avec les variables OA, EA et OU. L'axe 1 oppose les cantons de structure rurale prédominante (M3, K1 respectivement 80%, 70% d'agriculteurs en 1954) aux cantons ouvriers (G9, D8 respectivement 54 et 71% d'ouvriers en moyenne)

1 \ 2	CSP	EA	OA	AC	PL	CM	EM	OU	SE	CP
1	56	29	21	1	1	1	3	43	1	0
2	18	42	38	1	2	1	8	6	2	0
3	13	2	19	3	15	19	11	21	7	3
4	7	0	0	77	6	5	6	2	3	1
5	2	0	0	9	13	10	4	0	62	2

Tableau 7

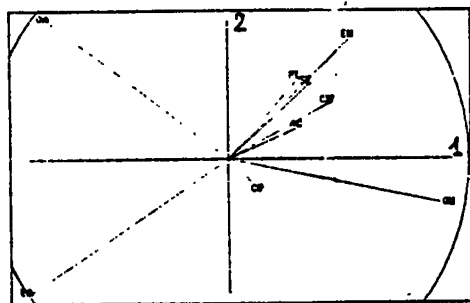


Figure 19

Cercle des corrélations locales 1-2

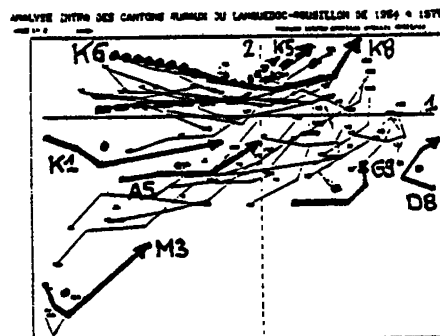


Figure 20

Plan de compromis 1-2

L'axe 2 sépare les cantons principalement selon leur type d'agriculture: familiale en bas (M3 - Lozère 86% d'exploitants dans la population rurale) industrielle en haut (K6 , Hérault, 39% d'exploitants parmi les ruraux). Si les catégories tertiaires interviennent assez peu sur cet axe (13% de  $var_U(F^2)$ ), elles apparaissent cependant corrélées négativement avec la catégorie EA ( surtout EM:  $r_U(EM, EA) = -0,63$ ). Et les cantons de structure tertiaire apparaissent en haut et à droite de ce plan.

L'axe 3 (cf. figures 21 et 22) oppose les cantons ouvriers des secteurs agricoles ( C9 Narbonne 43% de OA en 54, H4 Agde 37% en 54 et 27% en moyenne) et urbaines (D8, G9 cantons du Gard déjà cités) aux cantons tertiaires.

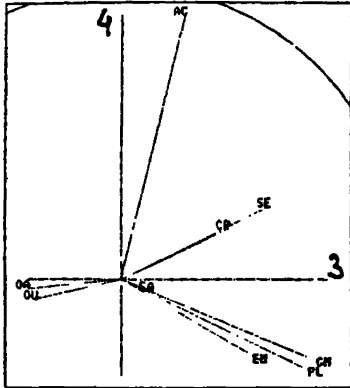


Figure 21

Cercle des corrélations locales

1-2

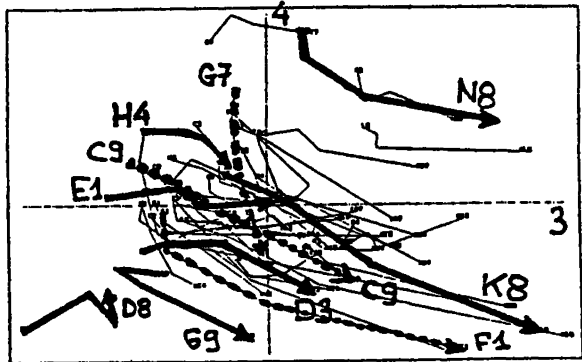


Figure 22

Plan des compromis 3-4

La visualisation des trajectoires sur le plan 1-2 montre une relative stabilité des cantons en projection sur ce plan. Aussi les corrélations totales et locales des catégories avec les facteurs sont proches. On en déduit que d'une manière générale les cantons gardent les mêmes positions relatives en projection sur le plan de compromis. Sur le plan 3-4, la situation est différente; des trajectoires traversent largement le plan de compromis. Les corrélations locales et totales divergent (on a par exemple  $r(EA, F3) = -0,53$ ) Au sens des corrélations totales ( donc en particulier pour comparer deux états successifs d'un même canton) l'axe 3 devient un axe d'opposition rural-tertiaire.

### 3- CONCLUSION:

Nous avons essayé sur ce jeu de données d'autres analyses factorielles locales, et même l'analyse en composantes principales classique. Elles ont toutes données d'assez bons résultats dans le sens que toutes permettaient de mettre en évidence les traits majeurs de l'évolution des cantons ou ceux de leur structure interne. Mais selon les méthodes choisies, les

contributions des catégories de grands effectifs étaient sans commune mesure avec leur fréquence dans la population. D'autre part, le mélange de la variance inter et de la variance intra, que réalise l'ACP classique, ne permet d'introduire la notion d'évolution dans l'interprétation qu'en visualisant les trajectoires à postériori, sur le plan de représentation. C'est pourquoi l'analyse locale, par la souplesse des choix qu'elle permet tant au niveau du choix du graphe que du poids des couples ou des variables nous semble bien adapté à ce type de données.

Il lui manque encore des outils de classification adaptés qui permettent de définir des classes de trajectoires homogènes, pour en simplifier l'interprétation. De tels outils, ainsi qu'un bon logiciel interactif graphique, devraient dans un proche avenir faciliter la tâche de ceux qui sont confrontés à ce genre de données.

BIBLIOGRAPHIE

---

- 1 CAILLES F. et PAGES J.P, Introduction à l'analyse des données, Société de Mathématiques Appliquées aux Sciences Humaines, 1976.
- 2 CARLIER A, "Analyse des évolutions sur table de contingence: quelques aspects opérationnels" , 1985 , article soumis au Congrès Inria, Analyse des Données et Informatique
- 3 LEBART L., Analyse Statistique de la contiguïté, 1966 , Thèse de troisième cycle, Paris.
- 4 LEBART L. et TABARD N., "Description statistique de certaines relations linéaires (Analyse des correspondances locales), 1971 , Rapport Cordes 13, 1973, Rapport Cordes Credoc.
- 5 LE FOLL Y., "L'Analyse des évolutions", 1972 , Actes de Petit Séminaire sur l'Analyse Factorielle, n° 2, p.22-26
- 6 LE FOLL Y., "Analyse en différences essentielles; méthode d'approche des processus multidimensionnels", 1973, Publication interne du Laboratoire d'Hydrologie Mathématique de l'Ecole des Mines de Paris.
- 7 LE FOLL Y., "Pondération des distances en Analyse Factorielle", 1982 , Statistiques et Analyse des Données, n° 7, 1, p.13-31.