Statistique et analyse des données

Alain Boudou

Analyse en composantes principales partielle

Statistique et analyse des données, tome 7, n° 2 (1982), p. 1-21

<http://www.numdam.org/item?id=SAD_1982__7_2_1_0>

Statistiques et Analyse de données
1982 - Vol. 7 n° 2 pp. 1-21

ANALYSE EN COMPOSANTES PRINCIPALES PARTIELLE

Alain BOUDOU

Laboratoire de Statistique et Probabilités
C.N.R.S.- E.R.A. 591
Université Paul Sabatier - Toulouse

Résumé : *Nous donnons une interprétation en termes de minimisation de l'inertie et en termes de minimisation de distances entre variables aléatoires de l'analyse de covariances partielles. La définition ainsi obtenue permet de faire la synthèse de méthodes apparemment fort différentes. Nous abordons les problèmes de représentations liés à ce type d'analyse. Parmi les différentes utilisations possibles de cette méthode, nous nous intéressons tout particulièrement au cas des séries chronologiques.*

Abstract : *This paper gives an interpretation of partial covariance analysis using two different minimisations : on one hand random variables distances, on the other hand inertia. The so obtained definition fits up a synthesis of a priori very different methods. Representation problems in touch with this kind of analysis are developped. Time series are specially focused among the different possible uses of the presented method.*

Mots clés : *analyse des données - analyse en composantes principales - séries chronologiques - analyse de covariances partielles.*

0 - INTRODUCTION

L'analyse en composantes principales d'une variable aléatoire est la recherche du sous-espace affine, de dimension donnée, qui minimise l'inertie. Le vecteur origine du sous-espace affine, qui est le centre de gravité du nuage des individus, peut être considéré comme une première approche de ceux-ci, approche qui est identique pour tous les individus. Afin de prendre en compte une éventuelle information supplémentaire concernant les individus, il peut être intéressant que cette approche soit fonction des individus. C'est-à-dire que l'origine du sous-espace affine sur lequel un individu donné est projeté, dans un but de représentation, dépende de celui-ci. Ainsi, dans le cas particulier où "l'information supplémentaire" serait fournie par une variable qualitative, l'origine serait fonction de la modalité prise par l'individu. L'objet de cet exposé est l'étude d'une analyse basée sur cette idée.

Nous la définissons en nous appuyant sur un exemple introductif pour lequel une telle analyse s'impose d'une façon naturelle, afin de tenir compte d'effets de niveaux. La méthode obtenue permet d'éliminer l'influence d'un éventuel phénomène exogène, ce qui, mathématiquement, se traduit par le fait que la fonction des individus "origine du sous-espace affine" doit appartenir à un sous-espace de variables aléatoires donné. Elle apparaît comme une analyse des résidus, sa réalisation se ramenant à une analyse des covariances partielles ce qui permet de donner à cette dernière une interprétation en termes de minimisation de l'inertie qui nous paraît originale.

Tout comme cela peut être fait pour l'analyse en composantes principales, l'analyse proposée peut être définie en termes de minimisation de distances entre variables aléatoires.

Nous abordons également les problèmes de représentation liés à ce type d'analyse.

Parmi les exemples d'applications considérés, le cas des fonctions aléatoires est important. Comme cas particulier simple, on retrouve une méthode d'analyse des séries chronologiques due à J. Obadia, B. Prieuret et M. Tenenhaus.

1.1 - Exemple introductif

On se propose d'analyser les résultats électoraux des 8 départements du Sud-Ouest aux élections présidentielles et législatives de 1981. L'espace de représentation des individus est $\mathbb{R}^4$, muni de la métique usuelle, la première (resp. la deuxième ; resp. la troisième, resp. la quatrième) coordonnée correspond au pourcentage de

voix communistes (resp. de la gauche non communiste ; resp. de la droite ; resp. qui se sont abstenues). Les données ont été calculées à partir des résultats fournis par le quotidien "Le Monde" du 16 juin 1981. Ces calculs, par les regroupements qu'ils ont nécessités, peuvent, bien entendu être contestés; aussi accorderons-nous un crédit circonspect aux résultats. Nous pouvons envisager trois analyses en composantes principales (A.C.P.) centrées suivant que l'on considère les résultats des présidentielles, des législatives ou bien des présidentielles et des législatives. Pour chacune de ces trois A.C.P., le poids des individus est proportionnel au nombre d'électeurs du département correspondant.

Le tableau II et les planches I et II résument les deux premières étapes de chacune des trois A.C.P. Le premier facteur de la première analyse et le premier facteur de la deuxième analyse sont voisins et correspondent à une contribution à la variance très élevée. Ils réalisent une dichotomie "droite-gauche" très classique pour cette sorte d'analyse (cf [1] par exemple). Les combinaisons linéaires correspondantes constituent des indices politiquement très significatifs. De plus, les représentations des départements obtenues à partir des plans principaux sont semblables. Quant aux résultats de la troisième analyse, contrairement à toute attente, ils sont très différents : la quatrième coordonnée du premier facteur est très élevée. Cela s'explique par le fait que les taux d'abstentions sensiblement équivalents pour un même type d'élection, sont très différents d'un type à l'autre. Ainsi, la variable "abstention" a-t-elle une variance faible pour les deux premières A.C.P. et beaucoup plus forte pour la troisième, ce qui se traduit dans l'expression du premier facteur par un coefficient correspondant plus élevé. Quoiqu'il en soit, cette troisièmeanalyse présente peu d'intérêts pour le politologue ; de plus, on peut regretter cette dissemblance entre analyse globale et analyses partielles. L'un des buts de l'analyse que nous étudions dans ce texte est de tenter de remédier à cet inconvénient. Pour cela, nous commençons par reformuler le problème précédent, ensuite nous l'étudions dans un cadre plus général.

Si l'on désigne par E_1 (resp.E_2) l'ensemble des individus " département-présidentielle"(resp."département-législative"), la première (resp.la deuxième ; resp. la troisième) analyse concerne les résultats de l'ensemble E_1 (resp. E_2 ; resp.$E=E_1 \cup E_2$). Notons V_1 (resp. V_2 ; resp.V) la matrice de variance-covariance associée à la première (resp. la deuxième ; resp. la troisième) A.C.P. et M_1 (resp. M_2 ; resp. M) le centre de gravité de l'ensemble correspondant des individus. Nous avons, avec des notations évidentes, d'après une relation bien connue liant la variance totale aux variances intra et inter, l'égalité matricielle :

$$V = \sum_{i=1}^{2} \frac{1}{2} V_i + \sum_{i=1}^{2} \frac{1}{2}(M - M_i)^t (M - M_i) .$$

Désignons par X l'application de E dans $\mathbb{R}^4$ qui à un individu e associe le point représentatif de celui-ci dans $\mathbb{R}^4$. Nous savons (cf[2]) que les k ($k \leq 4$) premières étapes de l'A.C.P. centrée de X sont, avec des notations évidentes, la

recherche d'un élément h de R^4 et d'un sous-espace U de dimension k rendant minimal $\sum_{e \in E} p_e \, d^2(X(e), h+U)$

On cherche donc un sous-espace affine (h+U) tel que la projection des points représentatifs des individus sur celui-ci donne une image de ceux-ci la plus "représentative possible" (on sait que h=M). D'après ce qui précède (premier facteur de la première A.C.P. et premier facteur de la seconde voisines, contributions correspondantes à la variance élevées), il est clair que V_1 et V_2 sont peu différentes, donc l'opposition entre les deux premières A.C.P. et la troisième, ou bien encore entre V_1, V_2 et V provient du terme $\sum_{i=1}^{2} \frac{1}{2}(M-M_i)^t(M-M_i)$ c'est-à-dire de la différence entre les centres de gravité M_1 et M_2. D'où l'idée de projeter les individus sur des sous-espaces affines d'origines différentes suivant qu'ils appartiennent à E_1 ou à E_2. C'est cette idée, a priori simple et classique pour le praticien, que nous proposons de développer en essayant d'y donner un fondement théorique. Dans le cas particulier ci-dessus, on est ainsi conduit (pour les k premières étapes) à la :

recherche de (h_1, h_2) de $\mathbb{R}^4 \times \mathbb{R}^4$ et de U sous-espace de dimension k rendant minimal $\sum_{e \in E_1} p_e \, d^2(X(e), h_1+U) + \sum_{e \in E_2} p_e \, d^2(X(e), h_2+U)$.

Si l'on note P la mesure de probabilité sur $(E, \mathcal{P}(E))$ induite par la pondération des individus, on peut écrire :

$$\sum_{e \in E_i} p_e \; d^2(X(e), h_i+U) = \int_{E_i} d^2(X(e), h_i+U) \, dP(e) \quad \text{(pour } i=1,2).$$

De plus, comme il y a correspondance biunivoque entre l'ensemble des couples (h_1, h_2) de $\mathbb{R}^4$ et l'espace vectoriel G des variables aléatoires (v.a) à valeurs dans $\mathbb{R}^4$, constantes sur chacun des sous-ensembles E_1 et E_2, la méthode précédente (pour les k premières étapes) est la

recherche d'un élément g de G et d'un sous-espace U de $\mathbb{R}^4$ de dimension k minimisant $\int_E d^2(X(e), g(e)+U) dP(e)$.

Ce dernier problème, que l'on résoud dans le paragraphe suivant et dont on montre que l'analyse qu'il définit répond à nos aspirations en ce qui concerne l'exemple introductif, se prête à diverses généralisations :

- au lieu du sous-espace G, ensemble de v.a. vectorielles dont les composantes sont des v.a. $\{\phi, E_1, E_2, E\}$-mesurables, on peut considérer un sous-espace de v.a.-vectorielles dont les composantes appartiennent à un sous-espace S donné de v.a. réelles (à titre d'exemples, visant des objectifs différents, S peut être le sous-espace engendré par une famille de v.a. ou bien par une famille d'indicatrices ou encore l'orthogonal à ce dernier),
- l'espace probabilisé de référence $(E, \mathcal{P}(E), P)$ peut être quelconque et, en particulier non nécessairement fini. Généralisation qui n'est pas un simple exercice de mathématique car la modélisation de phénomènes physiques fait souvent appel, d'une façon naturelle, à de tels espaces. Certes, même dans ce cas, la mise en oeuvre pratique de l'analyse nécessite de se cantonner au cas fini, mais alors cela nécessite une "approche" de l'analyse "théorique" et il faut bien que cette dernière soit définie d'une façon correcte afin de pouvoir appréhender convenablement le problème de convergence sous-jacent à l'approximation. Notons également que l'espace probabilisé peut être un espace produit, ce qui permet d'envisager le cas des fonctions aléatoires.
- enfin, on peut considérer des v.a. hilbertiennes ce qui englobe le cas particulier important où les v.a. sont à valeurs dans $\mathbb{R}^p$ muni d'une métrique euclidienne quelconque.

Les objectifs que nous nous sommes fixés et leur diversité nécessitent que nous nous plaçions dans un cadre quelque peu théorique. Nous nous apercevons que l'on obtient ainsi une formulation commune pour des problèmes divers, nous en examinerons certains au niveau des applications, nous retrouverons ainsi parfois des analyses connues mais dont les fondements théoriques n'étaient pas toujours, du moins à notre connaissance, très clairs.

2 - S-ANALYSE EN COMPOSANTES PRINCIPALES

2.1 - Notations et rappels

Tous les espaces de Hilbert considérés sont réels et séparables. Lorsque U est un sous-espace fermé quelconque d'un espace de Hilbert, on désigne par P_U le projecteur orthogonal sur U et par $U^\perp$ le supplémentaire orthogonal de U.

Dans la suite de ce paragraphe, $\mathcal{H}$ désigne un espace de Hilbert, muni de sa tribu borélienne, et η une mesure bornée définie sur l'espace mesurable $(E,\mathcal{F})$. Nous savons que l'espace de Hilbert $L^2(E,\mathcal{F},\eta)\otimes\mathcal{H}$ est identifiable à l'espace $\sigma_2(L^2(E,\mathcal{F},\eta),\mathcal{H})$ des opérateurs de Hilbert-Schmidt de $L^2(E,\mathcal{F},\eta)$ dans $\mathcal{H}$; nous ferons fréquemment par la suite cette identification.
Si X est un élément quelconque de $L^2_{\mathcal{H}}(\mathcal{F})$, l'application $\tilde{X} = \begin{cases} L^2(\mathcal{F}) \longrightarrow \mathcal{H} \\ y \longrightarrow \int_E y(e)X(e)d\eta(e) \end{cases}$

est un opérateur de Hilbert-Schmidt dont l'adjoint $\tilde{X}^*$ est défini par :
$(\tilde{X}^*h)(e) = \langle X(e),h\rangle$ (pour tout (h,e) de $\mathcal{H}\times E$).

L'application $\mathcal{J} = \begin{cases} L^2_{\mathcal{H}}(\mathcal{F}) \longrightarrow \sigma_2(L^2(\mathcal{F}),\mathcal{H}) \\ X \longrightarrow \tilde{X} \end{cases}$ est une isométrie .

Remarquons que, pour toute sous-tribu $\mathcal{B}$ de $\mathcal{F}$, on a : $\mathcal{J}(L^2_{\mathcal{H}}(\mathcal{B})) = \sigma_2(L^2(\mathcal{B}),\mathcal{H})$. Dans le cas particulier où $(E,\mathcal{F},\eta)$ est un espace probabilisé, l'A.C.P. non centrée d'unev.a. hilbertienne X est obtenue (cf [3])par l'analyse spectrale de l'un des opérateurs : $\tilde{X}$, $\tilde{X}^*$, $\tilde{X}\circ\tilde{X}^*$ ou $\tilde{X}^*\circ\tilde{X}$.

2.2. Définition de la S-analyse en composantes principales.

Soient X un élément de $L^2_{\mathcal{H}}(\mathcal{F})$ et S un sous-esapce fermé de $L^2(\mathcal{F})$. On note G l'image dans $L^2_{\mathcal{H}}(\mathcal{F})$ par $\mathcal{J}^{-1}$ du sous-espace fermé $S\otimes\mathcal{H}$ de $L^2(\mathcal{F})\otimes\mathcal{H}$. Il vient, avec les conventions de notations définies au § 2.1., la :

*
Proposition 1 : *Si δ est un élément de G et U un sous-espace de $\mathcal{H}$ de dimension k, l'application*

$\begin{cases} E \longrightarrow \mathbb{R} \\ e \longrightarrow d^2(X(e),\delta(e)+U) \end{cases}$ *est $\mathcal{F}$-mesurable et η-intégrable ; de plus, on a l'égalité :*

$$\int_E d^2(X(e),\delta(e)+U)d\eta(e) = \|X-P_G X\|^2_{L^2_{\mathcal{H}}(\mathcal{F})} - 2\langle \widetilde{X-P_G X}\circ\widetilde{X-P_G X}^*, P_U\rangle + \|P_{U^\perp}\circ(P_G X-\delta)\|^2_{L^2_{\mathcal{H}}(\mathcal{F})} .$$

---

* On peut se procurer la démonstration de cette proposition en s'adressant à l'auteur. Il en sera ainsi pour toute proposition qui dans la suite du texte sera précédée d'une astérisque.

Bien entendu, $d(\mathcal{X}(e), f(e)+U)$ est la distance, dans $\mathcal{H}$, de $\mathcal{X}(e)$ au sous-espace affine d'origine $f(e)$ et parallèle à U ; quant aux notations $\|\ \|_{L^2_{\mathcal{H}}(\mathcal{T})}$ et $<\ ,\ >_2$, elles désignent respectivement la norme de l'espace de Hilbert $L^2_{\mathcal{H}}(\mathcal{T})$ et le produit scalaire sur l'espace de Hilbert des opérateurs de Hilbert-Schmidt. Ce résultat nous permet de poser la :

Définition : *On appelle S-analyse en composantes principales de $\mathcal{X}$ l'analyse dont les k premières étapes sont la recherche d'un élément $\mathcal{G}$ de G et d'un sous-espace U de dimension k de $\mathcal{H}$ tels que la quantité $\int_E d^2(\mathcal{X}(e), \mathcal{G}(e)+U)d\eta(e)$ soit minimale.*

Notant $\sum_{i\in I} \lambda_i^2\, h_i \otimes h_i$ la décomposition de Schmidt de l'opérateur $\widetilde{\mathcal{X}-P_G\mathcal{X}} \circ \widetilde{\mathcal{X}-P_G\mathcal{X}}^*$ et U_k le sous-espace engendré par $h_1,\ldots,h_k$, nous savons (cf [1] p.14) que :

$$<\widetilde{\mathcal{X}-P_G\mathcal{X}} \circ \widetilde{\mathcal{X}-P_G\mathcal{X}}^*\ ,\ P_U>_2 \ \leq\ <\widetilde{\mathcal{X}-P_G\mathcal{X}} \circ \widetilde{\mathcal{X}-P_G\mathcal{X}}^*, P_{U_k}>_2\ ,$$

pour tout sous-espace U de $\mathcal{H}$ de dimension k.

f étant un élément quelconque de G, la proposition 1 permet d'écrire :

$$\int_E d^2(\mathcal{X}e, f(e)+U)d\eta(e) = \|\mathcal{X}-P_G\mathcal{X}\|^2 - <\widetilde{\mathcal{X}-P_G\mathcal{X}} \circ \widetilde{\mathcal{X}-P_G\mathcal{X}}^*, P_U>_2 + \|P_{U^\perp} \circ (P_G\mathcal{X}-f)\|^2,$$

$$\int_E d^2(\mathcal{X}e, (P_G\mathcal{X})(e)+U_k)d\eta(e) = \|\mathcal{X}-P_G\mathcal{X}\|^2 - <\widetilde{\mathcal{X}-P_G\mathcal{X}} \circ \widetilde{\mathcal{X}-P_G\mathcal{X}}^*, P_{U_k}>_2\ ,$$

ce qui, compte tenu de l'inégalité ci-dessus conduit à :

$$\int_E d^2(\mathcal{X}(e), (P_G\mathcal{X})(e)+U_k)d\eta(e) \leq \int_E d^2(\mathcal{X}(e), f(e)+U)d\eta(e),$$ il est donc clair que :

la réalisation des k premières étapes de la S-A.C.P. de $\mathcal{X}$ est obtenue à partir des k premières étapes de l'A.C.P. (non centrée) de $\mathcal{X}-P_G\mathcal{X}$; quant à l'élément optimal de G c'est $P_G\mathcal{X}$, projection de $\mathcal{X}$ sur G.

Le schéma de dualité associé à cette A.C.P. est le suivant :

$$\begin{array}{ccc} \mathcal{H} & \xleftarrow{\widetilde{P_{G^\perp}\mathcal{X}}} & L^2(\mathcal{T}) \\ I \downarrow \uparrow V & & W \downarrow \uparrow I \\ \mathcal{H} & \xrightarrow{\widetilde{P_{G^\perp}\mathcal{X}}^*} & L^2(\mathcal{T}) \end{array}$$

$$V = \widetilde{P_G\mathcal{X}} \circ \widetilde{P_G\mathcal{X}}^*$$

$$W = \widetilde{P_{G^\perp}\mathcal{Y}}^* \circ \widetilde{P_{G^\perp}\mathcal{Y}}$$

L'analyse que l'on propose est une généralisation de l'A.C.P. centrée : dans le cas particulier où S = $L^2(\{\phi,E\})$, tout élément de G est à valeurs constantes dans $\mathcal{H}$ et la S-A.C.P. est alors, d'après la définition, la recherche d'un sous-espace affine optimal. Ceci nous permet de voir la différence entre A.C.P. centrée et S-ACP (lorsque S est quelconque) d'une façon claire :

- l'origine du sous-espace affine sur lequel on projette les éléments du nuage est dans le premier cas indépendante de l'élément considéré,

- elle en est fonction dans le cas de la S-A.C.P..

Ainsi la S-A.C.P. apparaît comme une analyse en composantes principales qui permettrait d'exploiter une "information supplémentaire" relative aux individus, information dont l'expression mathématique serait le sous-espace S. Certes, l'analyse que l'on propose peut-être considérée comme une analyse résiduelle : l'A.C.P. de la v.a. $P_{G^\perp} X$, mais la définition que nous en donnons a le mérite de mettre en valeur le rôle joué par l'"information supplémentaire", c'est-à-dire par le sous-espace S ; il en est de même pour la nouvelle approche que nous en donnons au paragraphe suivant.

2.3 - Une autre approche de la S-A.C.P.

On sait (cf. [1])que les k premières étapes de l'A.C.P. (non centrée) d'une v.a. X reviennent à chercher l'opérateur de rang k le plus proche de $\tilde{X}$ dans $\sigma^2(L^2(\mathcal{F}),\mathcal{H})$, ou encore -utilisant l'isométrie entre cet espace et $L^2_{\mathcal{H}}(\mathcal{F})$- à chercher k couples (Z_i, W_i) (i=1,2,...,k) de $L^2(\mathcal{F}) \times \mathcal{H}$ de sorte que $\sum_{i=1}^{k} Z_i \otimes W_i$ soit le plus proche possible de X (au sens de la norme de $L^2_{\mathcal{H}}(\mathcal{F})$). Nous allons établir une propriété analogue pour la S-A.C.P.. Pour cela, désignant par H_1 et H_2 deux espaces de Hilbert, F un sous-espace fermé de H_1, nous utiliserons les résultats suivants :

*

Proposition 2 : *si T est un opérateur de Hilbert-Schmidt de H_1 dans H_2 et* $\sum_{i\in I} \mu_i e_i \otimes \delta_i$ *la décomposition de Schmidt de* $T \circ P_{F^\perp}$, *alors, pour tout V de* $\sigma_2(H_1,H_2)$ *et tout opérateur L de rang k de* H_1 *dans* H_2, *on a l'inégalité :*

$$\|T - (T \circ P_F + \sum_{i=1}^{k} \mu_i e_i \otimes \delta_i)\|_2 \leq \|T - (V \circ P_F + L)\|$$

*

Lemme : *Pour tout y de* $L^2_{\mathcal{H}}(\mathcal{F})$, *l'image* $\widetilde{P_G y}$ *de* $P_G y$ *dans* $\sigma_2(L^2(\mathcal{F}),\mathcal{H})$ *est l'opérateur* $\tilde{y} \circ P_S$.

Grâce en particulier au lemme, il vient :

$$\widetilde{\mathcal{X} - P_G \mathcal{X}} = \widetilde{\mathcal{X}} - \widetilde{P_G \mathcal{X}} = \widetilde{\mathcal{X}} - \widetilde{\mathcal{X}} \circ P_S = \widetilde{\mathcal{X}} \circ P_{S^\perp} .$$

La décomposition de Schmidt $\sum_{i \in I} \lambda_i x_i \otimes h_i$ de $\widetilde{\mathcal{X} - P_G \mathcal{X}}$ est donc aussi celle de $\widetilde{\mathcal{X}} \circ P_{S^\perp}$.

Pour tout k-uple $((Z_1,W_1),\ldots,(Z_k,W_k))$ d'éléments de $L^2(\mathcal{F}) \times \mathcal{H}$, $\sum_{i=1}^{k} Z_i \otimes W_i$ est un opérateur de rang k ; en utilisant la proposition 2, on a donc, pour toute v.a. hilbertienne y de G :

$$\| \widetilde{\mathcal{X}} - (\widetilde{\mathcal{X}} \circ P_S + \sum_{i=1}^{k} \lambda_i x_i \otimes h_i) \|_2 \leq \| \widetilde{\mathcal{X}} - (\widetilde{y} \circ P_S + \sum_{i=1}^{k} Z_i \otimes W_i) \|_2 , \text{ soit encore}$$

(cf. lemme) :

$$\| \widetilde{\mathcal{X}} - (\widetilde{P_G \mathcal{X}} + \sum_{i=1}^{k} \lambda_i x_i \otimes h_i \|_2 \leq \| \widetilde{\mathcal{X}} - (\widetilde{P_G y} + \sum_{i=1}^{k} Z_i \otimes W_i) \|_2 .$$

Comme $\mathcal{J}^{-1}(x_i \otimes h_i) = x_i h_i$ et $\mathcal{J}^{-1}(Z_i \otimes W_i) = Z_i W_i$, cette inégalité s'écrit dans $L^2_{\mathcal{H}}(\mathcal{F})$:

$$\| \mathcal{X} - (P_G \mathcal{X} + \sum_{i=1}^{k} \lambda_i x_i h_i \| \leq \| \mathcal{X} - (P_G y + \sum_{i=1}^{k} Z_i W_i) \| = \| \mathcal{X} - (y + \sum_{i=1}^{k} Z_i W_i) \|$$

d'où la ::

Proposition 3 : *Les k premières étapes de la S-A.C.P. de la v.a. $\mathcal{X}$ peuvent être définies comme la recherche d'une v.a. y de G et de k éléments $(Z_1,W_1),\ldots,(Z_k,W_k)$ de $L^2(\mathcal{F}) \times \mathcal{H}$ tels que $y + \sum_{i=1}^{k} Z_i W_i$ soit le plus proche possible de $\mathcal{X}$ dans $L^2_{\mathcal{H}}(\mathcal{F})$.*

Cette autre approche, surtout intéressante pour l'analyse de fonctions aléatoires, permet de mieux cerner le rôle joué par le sous-espace G. De plus, elle met en évidence la décomposition obtenue ; en effet, nous avons :

$\widetilde{\mathcal{X}} = \widetilde{P_G \mathcal{X}} + \sum_{i \in I} \lambda_i x_i \otimes h_i$; soit encore dans $L^2_{\mathcal{H}}(\mathcal{F})$:

$\mathcal{X} = P_G \mathcal{X} + \sum_{i \in I} \lambda_i x_i h_i$.

Du fait que $\sum_{i \in I} \lambda_i x_i \otimes h_i$ est la décomposition de Schmidt de $\widetilde{\mathcal{X}} \circ P_{S^\perp}$, les composantes principales x_i sont orthogonales au sous-espace S. Ainsi, lorsque 1_E appartient à S, les composantes principales x_i sont centrées et sans corrélation avec un élément quelconque de S : elles donnent une description de $\mathcal{X}$ "dégagée" de l'in-

fluence due au phénomène sous-jacent à S, ce qui est bien confirmé par la remarque ci-dessous.

Remarque : si $S = L^2(\mathcal{D})$ où $\mathcal{D}$ est une sous-tribu indépendante de $\mathcal{X}$, alors $G = L^2_{\mathcal{H}}(\mathcal{D})$ (cf. § 2.1.). La projection de $\mathcal{X}$ sur G est $E^{\mathcal{D}}\mathcal{X}$, espérance conditionnelle de $\mathcal{X}$ par rapport à $\mathcal{D}$. Or, $\mathcal{D}$ étant indépendante de $\mathcal{Y}$, cette dernière est égale à $E\mathcal{X}$; ceci est une simple extension au cas desv.a. hilbertiennes d'un résultat bien connu (cf. [5] p.138) pour les v.a. réelles. Il est alors clair que l'A.C.P. centrée est identique à la S-A.C.P., ce qui paraît naturel car celle-ci, d'après ce qui précède, analyse ce qui est indépendant du phénomène sous-jacent à S, donc ici à $\mathcal{D}$, phénomène dont l'influence est nulle d'après les hypothèses faites.

2.4. Pseudo-composantes principales et représentation des individus

Les notations sont celles du paragraphe précédent. Il paraît légitime de représenter l'individu e par le sous-ensemble $\{\lambda_i\, x_i(e)\}_{i\in I}$ de $\mathbb{R}$, mais alors, contrairement à ce qui se passe pour l'A.C.P. classique, les distances ne sont plus respectées c'est-à-dire que pour un couple (e,e') quelconque d'individus la quantité $\sum_{i\in I} (\lambda_i\, x_i(e) - \lambda_i\, x_i(e'))^2$ peut être différente du carré de la distance de $\mathcal{X}(e)$ à $\mathcal{X}(e')$. C'est pour pallier un tel défaut que nous introduisons la :

Définition : *nous appelons $i^{ème}$ pseudo-composante principale la v.a. réelle* $C_i = P_S \circ \tilde{\mathcal{X}}^* h_i + \lambda_i x_i$.

On montre facilement que, pour tout i de I, C_i est l'image de h_i par $\tilde{\mathcal{X}}^*$. Sans restreindre la généralité on peut supposer que $\{h_i\}_{i\in I}$ est une base orthonormée de $\mathcal{H}$ (certains λ_i étant éventuellement nuls) et dès lors (cf. [1])on a :

$\tilde{\mathcal{X}} = \sum_{i\in I} (\tilde{\mathcal{X}}^* h_i) \otimes h_i$ soit $\tilde{\mathcal{X}} = \sum_{i\in I} C_i \otimes h_i$, ce qui transposé dans $L^2_{\mathcal{H}}(\mathcal{F})$ donne :

$\mathcal{X} = \sum_{i\in I} C_i . h_i$ ou bien encore $\mathcal{X}(e) = \sum_{i\in I} C_i(e) . h_i$, pour tout e de E. Il est donc clair que la représentation à l'aide des pseudo-composantes principales conserve les distances en ce sens que, pour tout couple (e,e') d'individus, on a :

$$d^2(\mathcal{X}(e), \mathcal{X}(e')) = \sum_{i\in I} (C_i(e) - C_i(e'))^2 .$$

2.5. Le cas de l'exemple introductif

Revenons à l'exemple introductif et choisissons pour sous-espace S l'ensemble $L^2(\{\phi,E_1,E_2,E\})$ des v.a. vectorielles constantes sur chacun des E_i. Le sous-espace G est $L^2_{\mathbb{R}^4}(\{\phi,E_1,E_2,E\})$. La S-A.C.P. de X est bien, d'après sa définition, l'analyse que nous avions projetée de réaliser au paragraphe 1. Sa réalisation, comme en témoignent le tableau II et la planche III, répond bien à nos souhaits. Notons que la représentation à partir des pseudo-composantes principales (cf. planche IV) donne une idée correcte de l'évolution des départements d'un type d'élection à l'autre : forte poussée abstentionniste .

3 - LES APPLICATIONS

L'objet de ce paragraphe est l'étude de quelques exemples d'applications de la S-A.C.P., c'est-à-dire l'examen des analyses obtenues lorsqu'on effectue tel ou tel choix pour le sous-espace S. Ces applications peuvent se classer en deux grandes catégories selon la nature de l'espace mesuré de référence : cas des variables aléatoires ou cas des fonctions aléatoires (espace produit).

3.1. Cas des variables aléatoires

Dans ce paragraphe X est une v.a. définie sur l'espace probabilisé $(\Omega,\mathcal{A},P)$ à valeurs dans un espace de Hilbert H et de norme carrée P-intégrable.

3.1.1. *Analyse des covariances partielles*

Nous désirons analyser X après l'élimination de l'influence des "variables exogènes" $Z_1,\ldots,Z_q$, qui sont des v.a. réelles de carré P-intégrable. Pour cela choisissant $S = \text{vect}\{Z_1,\ldots,Z_q\}$, sous-espace de $L^2(\mathcal{A})$ engendré par $Z_1,\ldots,Z_q$, nous proposons d'effectuer la S-A.C.P. de la v.a. X. Notant Z l'application

$$\begin{cases} \Omega \longrightarrow \mathbb{R}^q \\ \omega \longrightarrow (Z_1(\omega),\ldots,Z_q(\omega)) \end{cases}$$, qui de toute évidence est une v.a. à valeurs dans l'espace de Hilbert $\mathbb{R}^q$ (muni de la métrique usuelle) de norme carrée P-intégrable, il vient la

*

Proposition 4 : *Le sous-espace G, image de $S \otimes H$ dans $L^2_H(\mathcal{A})$, est l'ensemble des éléments pouvant se mettre sous la forme $\varphi \circ Z$ où φ est une application linéaire de $\mathbb{R}^q$ dans H.*

Et donc les k premières étapes de la S-A.C.P. sont la

recherche de l'application linéaire φ *et du sous-espace* U *de dimension* k *de* H *minimisant* $\int_\Omega d^2(X(\omega), \varphi((Z_1(\omega),\ldots,Z_q(\omega))) + U)\,dP(\omega)$.

Cette analyse est l'A.C.P. de la v.a. $X-P_G X$ dont l'opérateur de covariance, avec les notations que l'on a adoptées au § 2.1. et compte tenu du lemme du § 2.3., est :

$$V = \widetilde{X - P_G X} \circ \widetilde{X - P_G X}^* = (\tilde{X} - \tilde{X} \circ P_S) \circ (\tilde{X}^* - P_S \circ \tilde{X}^*) = \tilde{X} \circ P_{S^\perp} \circ \tilde{X}^*.$$

L'opérateur $P_{S^\perp}$, étant l'opérateur qui effectue la projection de tout élément de $L^2(\mathcal{G})$ sur le sous-espace orthogonal au sous-espace engendré par $Z_1,\ldots,Z_q$, il est clair que l'on obtient l'analyse des covariances partielles telle que celle-ci est définie en [4] page 300. Le fait de considérer cette analyse comme une S-A.C.P. permet de l'interpréter en termes de minimisation de l'inertie, ce qui à notre connaissance n'a pas été fait jusqu'à présent. Cela permet également de mieux cerner le rôle joué par les variables exogènes et éventuellement d'envisager quelques variantes. Ainsi, dans l'hypothèse où la linéarité traduit mal l'effet des variables exogènes sur X, on peut choisir pour sous-espace S l'espace $L^2(\mathcal{B})$ où $\mathcal{B}$ est la sous-tribu engendrée par Z ; G est alors $L^2_H(\mathcal{B})$ et est isométrique à $L^2_H(\mathbb{R}^q, \mathcal{B}_{\mathbb{R}^q}, P_Z)$, $\mathcal{B}_{\mathbb{R}^q}$ étant la tribu borélienne de $\mathbb{R}^q$ et P_Z la probabilité image de P par Z, par l'application $\begin{cases} L^2_H(\mathcal{B}_{\mathbb{R}^q}) \longrightarrow L^2_H(\mathcal{B}) \\ \varphi \longrightarrow \varphi \circ Z \end{cases}$ et donc les k premières étapes de la S-A.C.P. sont la

recherche de l'élément φ *de* $L^2_H(\mathcal{B}_{\mathbb{R}^q})$ *et du sous-espace* U *de dimension* k *de* H *minimisant* $\int_\Omega d^2(X(\omega), \varphi((Z_1(\omega),\ldots,Z_q(\omega))) + U)\,dP(\omega)$.

<u>Remarque</u> : dans le cas particulier où $Z_1 = 1_{B_1},\ldots,Z_q = 1_{B_q}$, $\{B_1,\ldots,B_q\}$ étant une partition de Ω, $\mathcal{B}$, sous-tribu engendrée par Z, est également la sous-tribu engendrée par la partition $\{B_1,\ldots,B_q\}$. Dès lors, $L^2(\mathcal{B})$ est le sous-espace engendré par $Z_1,\ldots,Z_q$ et donc les deux analyses précédentes sont confondues. On obtient une analyse qui permet de prendre en compte une partition de la population, ce qui dans la pratique est parfois une exigence très naturelle. Les composantes principales fournissent une description des différents lots dans un sous-espace commun. Si l'on désigne, pour tout i de $\{1,\ldots,q\}$, par V_i l'opérateur de cova-

riance associé à l'A.C.P. de $X_{|B_i}$, analyse partielle qui prend en considération uniquement les individus de B_i, un simple calcul nous fournit les égalités :

$$V_c = \sum_{i=1}^{q} V_i + \sum_{i=1}^{q} P(B_i) \left\{ EX - \frac{1}{P(B_i)} \int_{B_i} X\,dP \right\} \otimes \left\{ EX - \frac{1}{P(B_i)} \int_{B_i} X dP \right\} \text{ et } V = \sum_{i=1}^{q} V_i,$$

où V_c est l'opérateur de covariance de la v.a. $X - EX$. Ces relations traduisent les liens existant entre les facteurs principaux des différentes analyses possibles, la première lie d'une façon classique la variance totale aux variances entra et inter. L'exemple introductif est une application d'une telle S-A.C.P. .

3.1.2. *Analyse sous-contraintes.*

Soit $\mathcal{B}$ une sous-tribu quelconque de $\mathcal{A}$. En [1] est étudiée une "analyse en composantes principales sous contrainte de $\mathcal{B}$ -mesurabilité" qui se ramène à l'A.C.P. de la v.a. $E^{\mathcal{B}} X$. La S-A.C.P. permet de retrouver cette analyse lui donnant ainsi une nouvelle interprétation.

En effet, choisissons pour sous-espace S l'orthogonal de $L^2(\mathcal{B})$, on vérifie facilement que la S-A.C.P. de X nécessite l'A.C.P. de la v.a. $E^{\mathcal{B}} X$, c'est-à-dire l'A.C.P. sous contrainte de $\mathcal{B}$ -mesurabilité de X ; les k premières étapes de celle-ci sont donc la

recherche de la v.a. hilbertienne ϕ de norme carrée P-intégrable, d'espérance conditionnelle par rapport à $\mathcal{B}$ nulle, et du sous-espace U de H de dimension k minimisant $\int_\Omega d^2(X(\omega), \phi(\omega) + U)\,dP(\omega)$.

3.2. - Cas de fonctions aléatoires.

Soient (T,ξ,μ) et $(\Omega,\mathcal{A},P)$ deux espaces probabilisés et H un espace de Hilbert. Une fonction aléatoire hilbertienne (f.a.h.) est un élément quelconque de $L^2_H(\mathcal{A} \otimes \xi)$. Soit X un tel élément, il nous arrivera dans la suite de cet exposé de désigner la f.a.h. X par $(X_t)_{t\in T}$, où X_t est l'application partielle X(.,t) (qui appartient à $L^2_H(\mathcal{A})$). Une f.a.h. étant une application mesurable (définie sur $(\Omega\times T, \mathcal{A}\otimes\xi, P\otimes\mu)$) peut, du point de vue formel, être considérée comme une v.a. hilbertienne. Le but du § 3.2. est l'étude de diverses S-A.C.P. que l'on peut ainsi obtenir. Ces diverses méthodes nous permettrons de résoudre certains problèmes liés au temps ; en outre le fait qu'elles soient des S-A.C.P. facilitera leur comparaison.

3.2.1. *La S-analyse en composantes principales pour résoudre un problème de centrage.*

Pour analyser la f.a.h. X, nous avons proposé en [1] de faire l'A.C.P. de la "variable aléatoire hilbertienne" $X = \begin{cases} \Omega\times T \to H \\ (\omega,t) \to X(\omega,t) \end{cases}$. En analyse des données, il est souvent intéressant de procéder à un centrage préalable des variables, ici nous avons deux possibilités : ou bien centrer la "variable aléatoire" X, ou bien centrer chacune des v.a. X_t. Nous allons voir comment la S-A.C.P. apporte un élément de réponse à ce problème de choix.

Prenant $S = L^2(\{\phi,\Omega\} \otimes \xi)$, G est donc $L^2_H(\{\phi,\Omega\} \otimes \xi)$ et la projection de X sur G est l'application $\begin{cases} \Omega\times T \to H \\ (\omega,t) \to EX_t \end{cases}$. Le S-A.C.P. de X est alors l'A.C.P. de la "variable aléatoire" $\begin{cases} \Omega\times T \to H \\ (\omega,t) \to X(\omega,t) - EX_t \end{cases}$. On obtient ainsi le deuxième type de centrage indiqué ci-dessus.

D'après la définition de la S-A.C.P., les k premières étapes de cette analyse sont la recherche de l'élément h de G et du sous-espace U, de dimension k, de H minimisant $\int_{\Omega\times T} d^2(X_t(\omega),g(\omega,t)+U)\,dP\otimes\mu(\omega,t)$. L'élément h de G étant $\{\phi,\Omega\}\otimes\xi$-mesurable, l'application partielle h(.,t) est $\{\phi,\Omega\}$-mesurable c'est-à-dire qu'elle prend une valeur constante que l'on notera f(t) ; quant à h(ω,.) elle est ξ-mesurable. Par suite l'application $f = \begin{cases} T \to H \\ t \to f(t) \end{cases}$ appartient à $L^2_H(\xi)$. Utilisant le théorème de Fubini, l'expression à minimiser peut se mettre sous la forme

$$\int_T \left(\int_\Omega d^2(X_t(\omega),f(t)+U)dP(\omega) \right) d\mu(t)$$ ou bien encore $\int_T I^{X_t}_{f(t)+U}\, d\mu(t)$, si l'on désigne l'inertie du nuage $X_t(\Omega)$ par rapport au sous-espace affine f(t)+U par $I^{X_t}_{f(t)+U}$.

Et donc :

les k premières étapes de l'analyse de la f.a.h. $(X_t)_{t\in T}$ proposée sont la recherche de f appartenant à $L^2_H(\xi)$ et du sous-espace U, de dimension k, de H minimisant $\int_T I^{X_t}_{f(t)+U}\, d\mu(t)$.

Dans le cas où T, représentant le temps, est une partie de $\mathbb{R}$ si l'on suppose que l'injection canonique de T dans $\mathbb{R}$ appartient à $L^2(\xi)$, le fait qu'une composante principale quelconque $(y_t)_{t\in T}$ (élément de $L^2(G\otimes\xi)$) soit orthogonale à S (cf.§ 2.3) implique qu'il y ait absence de corrélation entre $(Ey_t)_{t\in T}$ et le temps ; c'est-à-dire que : $\int_T (t-\bar{t})(Ey_t - \overline{Ey_t})\,d\mu(t) = 0$ (avec $\bar{t} = \int_T t\,d\mu(t)$ et $\overline{Ey_t} = \int_T Ey_t\,d\mu(t)$). Il semble

donc, dans la mesure où l'on souhaite que les facteurs soient "dégagés" de l'influence du temps, que le deuxième type de centrage est le plus intéressant.
On peut formuler d'une façon analogue à ce qui vient d'être fait l'analyse correspondant ou premier type de centrage : c'est l'A.C.P. centrée de la "variable aléatoire" X, donc la recherche, pour les k premières étapes, de l'élément ℓ de H et du sous-espace U, de dimension k, de H minimisant $\int_{\Omega\times T} d^2(X(\omega,t),\ell+U)dp\otimes\mu(\omega,t)$, soit encore $\int_T I^{X_t}_{\ell+U}\, d\mu(t)$. La différence entre les deux types de centrage apparaît alors : dans un cas l'origine du sous-espace affine sur lequel on projette les éléments du nuage $X(\Omega\times T)$ dépend du "temps" (si T représente le temps), dans l'autre cas elle en est indépendante.

Remarque : dans [1] est étudié un exemple pratique d'analyse d'une f.a.h.. C'est le centrage correspondant à la S-A.C.P. examiné dans ce paragraphe qui a été adopté avec la représentation associée aux pseudo-composantes principales. De tels choix, motivés alors par l'empirisme, trouvent ici un fondement théorique.

3.2.2. *S-analyse en composantes principales et tendance polynomiale.*
Nous inspirant des travaux de J. Obadia, B. Prieuret et M. Tenenhaus (cf.[6] et [7]), que l'on retrouve ici comme cas particulier, on peut souhaiter que la projection $P_G X$ de X sur G-terme "ignoré" lors de l'analyse- ait une expression polynomiale.
Moyennant un choix convenable du sous-espace S, nous allons voir qu'une telle analyse entre dans le cadre proposé.
Nous supposons que T est une partie de $\mathbb{R}$ et que, pour tout j de $\{0,1,\ldots,n\}$, l'application $\begin{cases} T \longrightarrow \mathbb{R} \\ k \longrightarrow t^j \end{cases}$ est un élément de $L^2(\xi)$, ce qui entraine que

$\varphi_j = \begin{cases} \Omega\times T \longrightarrow \mathbb{R} \\ (\omega,t) \longrightarrow t^j \end{cases}$ appartient à $L^2(\{\phi\ \Omega\}\otimes\xi)$.

Soit S le sous-espace de $L^2(G\otimes\xi)$ engendré par $\varphi_0, \varphi_1,\ldots,\varphi_n$. Il vient alors la :

*
Proposition 5 : *Le sous-espace G, image dans $L^2_H(G\otimes\xi)$ de $S\otimes H$, est l'ensemble des f.a.h. du type* $\begin{cases} \Omega\times T \longrightarrow H \\ (\omega,t) \longrightarrow \sum_{j=0}^{n} t^j a_j \end{cases}$ *où* $a_0, a_1,\ldots,a_n$ *sont des éléments quelconques de H.*

Les k premières étapes de la S-A.C.P. de la "variable aléatoire" X sont donc la

recherche de n+1 éléments $a_0, a_1, \ldots, a_n$ *de H et du sous-espace U, de dimension k, de H minimisant* $\int_{\Omega \times T} d^2\left(X(\omega,t), \sum_{j=0}^{n} t^j a_j + U\right) dP \otimes \mu(\omega,t)$.

Elles sont également, d'après la deuxième approche donnée de la S-A.C.P., la

recherche de n+1 éléments $a_0, \ldots, a_n$ *de H et de k éléments* $(y_1,h_1), \ldots, (y_k,h_k)$ *de* $L^2(\mathcal{G} \otimes \xi) \times H$ *de sorte que* $\left\{ \begin{array}{l} \Omega \times T \to H \\ (\omega,t) \to \sum_{j=0}^{n} t^j a_j + \sum_{i=1}^{k} y_i(\omega,t).h_i \end{array} \right.$ *soit, au sens de la norme de* $L^2_H(\mathcal{G} \otimes \xi)$, *le plus proche possible de X.*

Cette dernière formulation met en relief le rôle joué par la tendance polynomiale (c'est-à-dire par G). Comme S est un sous-espace de $L^2(\{\phi \times \Omega\} \otimes \xi)$ (les φ_j appartenant à $L^2(\{\phi,\Omega\} \otimes \xi)$) S$\otimes$H est un sous-espace de $L^2(\{\phi\ \Omega\} \otimes \xi) \otimes H$ et G de $L^2_H(\{\phi,\Omega\} \otimes \xi)$. Par suite $P_G X$ peut être considéré comme la projection sur G de $\left\{ \begin{array}{l} \Omega \times T \to H \\ (\omega,t) \to EX_t \end{array} \right.$, projection de X sur $L^2_H(\phi,\Omega\} \otimes \xi)$. Vu la forme des éléments de G chercher la projection de X sur G revient à chercher n+1 éléments $a_0, a_1, \ldots, a_n$ de H minimisant $\int_{\Omega \times T} \| EX_t - \sum_{j=0}^{n} t^j a_j \|^2_{dP} \otimes \mu(\omega,t) = \int_T \| EX_t - \sum_{j=0}^{n} t^j a_j \|^2 \, d\mu(t)$. Ce résultat met en évidence le rôle de "résumé en moyenne" de la tendance polynomiale.

3.2.3. *Liens avec l'analyse proposée par J. Obadia, B. Prieuret et M. Tenenhaus*

Examinons maintenant la méthode d'analyse multidimendionnelle de séries chronologiques numériques proposée par J. Obadia, B. Prieuret et M. Tenenhaus (cf. [6] et [7]). Les ensembles Ω et T sont finis, de cardinaux respectifs m et q, et probabilisés par l'équiprobabilité. $X_1, \ldots, X_p$ désignent p f.a. réelles et X (resp. Y) la f.a. h. à valeurs dans $\mathbb{R}^p$ (resp. $\mathbb{R}^{p+1}$), muni du produit scalaire habituel, définie par : $X(\omega,t) = (X_1(\omega,t), \ldots, X_p(\omega,t))$ (resp. $Y(\omega,t) = (X_1(\omega,t), \ldots, X_p(\omega,t), t)$). Dans la suite de ce paragraphe, si v est un élément quelconque $(v_1, \ldots, v_p)$ de $\mathbb{R}^p$, nous convenons de noter (v, v_{p+1}) l'élément $(v_1, \ldots v_p, v_{p+1})$ de $\mathbb{R}^{p+1}$. Notons H_t l'ensemble $\{v' = (v,t)\ ;\ v \in \mathbb{R}^p\}$; H_o est donc le sous-espace engendré par les p premiers vecteurs de la base canonique de $\mathbb{R}^{p+1}$, quant à H_t c'est encore le sous-espace affine $(0, \ldots, 0, t) + H_o$.

La méthode d'analyse des p séries chronologiques $X_1, \ldots, X_p$ proposée par J. Obadia et M. Tenenhaus peut se formuler de la façon suivante :

Les k ($k \leq p$) premières étapes sont la recherche

- *de $n+1$ éléments $a'_0,\ldots,a'_n$ de $\mathbb{R}^{p+1}$ tels que pour tout t de T, $\sum_{j=0}^{n} t^j a'_j$ appartienne à H_t ,*
- *et de k éléments $(y_1,u'_1)\ldots(y_k,u'_k)$ de $L^2(\Omega\times T)\times\mathbb{R}^{p+1}$, $\{u'_1,\ldots,u'_k\}$ étant un système orthonormé de H_0, tels que*

$$\sum_{(\omega,t)\in\Omega\times T} \Big\| Y(\omega,t) - \sum_{j=0}^{n} t^j a'_j - \sum_{i=1}^{k} y_i(\omega,t)\, u'_i \Big\|^2_{\mathbb{R}^{p+1}} \text{ soit minimal.}$$

Notant, avec les conventions d'écriture adoptées, $a'_j = (a_j,\alpha_j)$ et $u'_i = (u_i,\beta_i)$ (a_i et u_i appartenant à $\mathbb{R}^p$), nous pouvons écrire l'expression à minimiser sous la forme

$$\sum_{(\omega,t)\in\Omega\times T} \frac{1}{mq} \Big\| \Big(X(\omega,t),t\Big) - \Big(\sum_{j=0}^{n} t^j a_j\,,\ \sum_{j=0}^{n} t^j \alpha_j\Big) - \Big(\sum_{i=1}^{k} y_i(\omega,t)u_i\,,\ \sum_{i=1}^{k} y_i(\omega,t)\beta_i\Big) \Big\|^2,$$

ou bien, puisque $\sum_{j=0}^{n} t^j a'_j$ (resp.u'_i) appartient à H_t (resp.H_0) ,

$$\sum_{(\omega,t)\in\Omega\times T} \frac{1}{mq} \Big\| X(\omega,t) - \sum_{j=0}^{n} t^j a_j - \sum_{i=1}^{n} y_i(\omega,t)u_i \Big\|^2 .$$

Il est alors clair que les k premières étapes sont la

recherche de $(n+1)$ éléments $a_0, a_1,\ldots,a_n$ de $\mathbb{R}^p$, de k éléments $y_1,\ldots,y_k$, de $L^2(\Omega\times T)$ et d'un système orthonormé $\{\mu_1,\ldots,u_k\}$ de $\mathbb{R}^p$ tels que

$$\begin{cases} \Omega\times T \to \mathbb{R}^p \\ (\omega,t) \longrightarrow \sum_{j=0}^{n} t^j a_j + \sum_{i=0}^{n} y_i(\omega,t)u_i \end{cases} \text{ soit le plus proche possible de } X \text{ dans } L^2_{\mathbb{R}^p}(\Omega\times T) .$$

On retrouve donc la méthode d'analyse proposée au § 3.2.2. à l'orthonormalité près des u_i , condition que l'on peut de fait omettre puisque les vecteurs optimaux obtenus vérifient cette condition. La possibilité de considérer l'analyse de J. Obadia et M. Tenenhaus comme une S-A.C.P. permet d'en donner une autre interprétation,plus simple du fait que l'introduction de la $p+1^{\text{ième}}$ composante, qui peut paraître un peu artificielle, s'avère inutile. Les développements mathématiques s'en trouvent simplifiés. Par ailleurs, il devient plus facile de faire la comparaison avec des méthodes voisines.

Conclusion

La S-A.C.P. permet de se dégager de l'influence d'un phénomène exogène ce que montre bien la remarque du § 2.3.. Il est clair qu'elle se ramène à une analyse des résidus, pratique connue depuis fort longtemps comme en témoignent certaines des applications envisagées, mais qui pouvait souvent apparaître comme un a priori dont les raisons étaient quelque peu pragmatiques. Cette étude justifie l'analyse des résidus ; de même que, dans le cas de l'A.C.P. classique, c'est la recherche d'un sous-espace affine qui justifie le centrage, pratique qui sans cela pourrait paraître artificielle.

La S-A.C.P. permet d'embrasser des cas apparemment très différents. On peut de la sorte donner une définition semblable à l'analyse sur matrice de covariances partielles et à l'analyse sous contraintes linéaires (cf. § 3.2.1.).

Le formalisme adopté permet de resituer ces analyses dans le cadre des travaux concernant les analyses factorielles les plus récents. Il parait être un outil indispensable pour aborder les problèmes de convergence ou lorsque, dans le cas de l'analyse des covariances partielles, on souhaite se dégager de la linéarité.

Les applications considérées ne sont pas exhaustives. En particulier en ce qui concerne les séries chronologiques on peut obtenir une analyse "désaisonnalisée" si l'on choisit un sous-espace S convenable.

REFERENCES

[1] A. BOUDOU : *Différents types d'analyses en composantes principales de fonctions aléatoires hilbertiennes ; étude de certaines contraintes.* Publication du Laboratoire de Statistique et Probabilités - N° 06-79, Université Paul Sabatier, Toulouse, 1979.

[2] F. CAILLEZ et J.P. PAGES : *Introduction à l'analyse des données.* Smash (1976).

[3] J. DAUXOIS et A. POUSSE : *Les analyses factorielles en calcul des Probabilités et en Statistique : essai d'étude synthétique.* Thèse, Université Paul Sabatier, Toulouse, 1976.

[4] L. LEBART, A. MORINEAU et J.P. FENELON : *Traitement des données statistiques.* Dunod, 1979.

[5] M. METIVIER : *Notions fondamentales de la théorie des probabilités.* Dunod, 1972.

[6] J. OBADIA et M. TENENHAUS : *Analyse multidimensionnelle de séries chronologiques nominales, ordinales ou numériques.* Journées internationales "Analyse des données et Informatique" - I.R.I.A. Versailles 7-9 septembre 1977.

[7] B. PRIEURET et M. TENENHAUS : *Analyse des séries chronologiques multidimensionnelles*, R.A.I.R.O., 1974.

[8] C.R. RAO : The use and interpretation of principal component analysis in applied research. Sankhya, ser.A,26,1964.

Je remercie les rapporteurs qui par leurs remarques constructives ont contribué à l'amélioration de la première rédaction de cet article.

Tableau I : données initiales en pourcentages de voix

	Ariège		Aveyron		Ht.Garonne		Gers		Pyr.Atl.		Pyr.Ht.		Tarn		Tarn & Gar.	
	Pres.	Legis.	Pres.	Legis.	Pres.	Legis.	Pres.	Legis.	Pres.	Legis.	Pres.	Legis.	Pres.	Legis.	Pres.	Legis.
Poids	0,027	0,027	0,055	0,055	0,137	0,137	0,035	0,035	0,102	0,102	0,044	0,044	0,064	0,064	0,036	0,036
Voix com.	16,25	15,41	7,38	6,01	12,26	8,84	11,17	8,61	8,54	6,41	14,87	12,74	12,02	8,12	11,21	8,06
Voix G. non C.	37,31	33,92	33,50	29,81	38,62	37,47	39,43	38,68	33,68	31,90	35,20	25,10	35,80	38,07	36,96	37,28
Voix de dr.	31,15	20,68	44,57	36,75	32,83	21,24	35,13	23,47	42,95	34,16	33,31	21,17	38,92	29,35	37,90	28,41
Abst.	15,29	30	14,10	27,43	16,30	32,45	14,27	29,23	14,84	34,16	16,61	31	13,26	24,46	13,93	26,64

Tableau II : les facteurs principaux

	Première analyse (résultats des pres.)		Deuxième analyse (résultats des légis.)		Troisième analyse (totalité des résultats)		S-Analyse en composantes principales	
	1er facteur	2ème facteur	1er facteur	2ème facteur	1er facteur	2ème facteur	1er facteur	2ème facteur
Communiste	0,37	-0,72	0,26	0,18	-0,02	-0,52	0,30	0,25
Gauche non com.	0,33	0,69	0,32	-0,76	0,05	-0,48	0,32	-0,76
Droite	-0,86	-0,05	-0,86	-0,02	-0,72	0,48	-0,87	-0,05
Abstention	0,13	-0,03	0,29	0,63	0,69	0,51	0,22	0,60
Contribution à la variance	0,90	0,05	0,82	0,11	0,80	0,17	0,85	0,09

Planche I : première analyse - plan principal

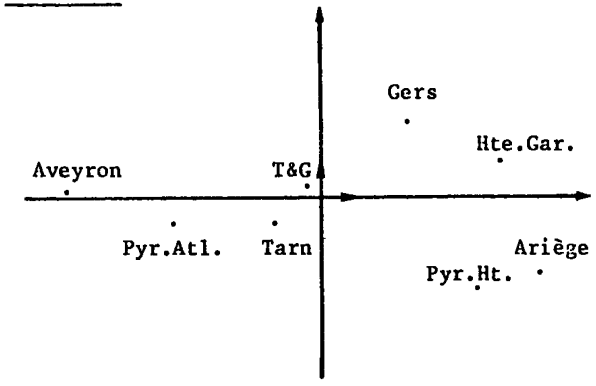

Planche II : deuxième analyse - plan principal

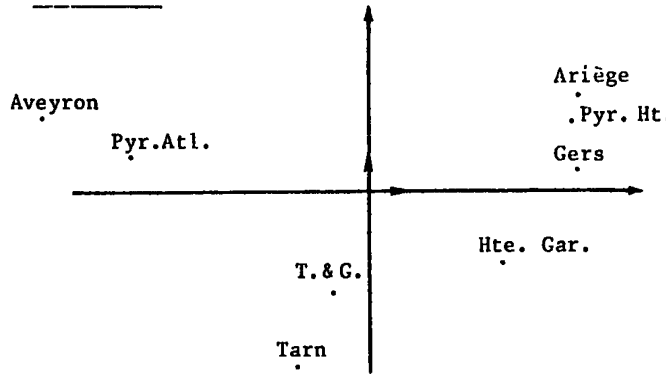

Planche III : S-A.C.P. (représentation à partir des cpst. pr.)
L : legislatives P : présidentielles

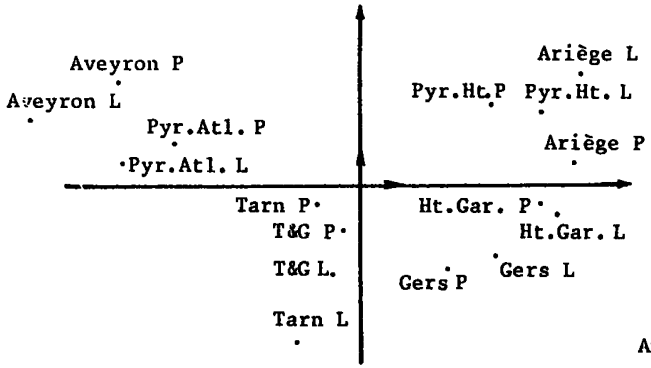

Planche IV : S-A.C.P. (représentation à partir des pseudo-cpst.pr.)