Statistique et analyse des données

Jacques Benasseni

Une amélioration d'un résultat concernant l'influence d'une unité statistique sur les valeurs propres en analyse en composantes principales

Statistique et analyse des données, tome 11, nº 1 (1986), p. 42-63

<http://www.numdam.org/item?id=SAD_1986__11_1_42_0>

Statistique et Analyse des Données
1986 - Vol.11 n°1 - pp. 42-63

UNE AMELIORATION D'UN RESULTAT CONCERNANT L'INFLUENCE D'UNE UNITE STATISTIQUE SUR LES VALEURS PROPRES EN ANALYSE EN COMPOSANTES PRINCIPALES

Jacques BENASSENI

Unité de Biométrie
ENSA-INRA-USTL
34060 MONTPELLIER CEDEX

Résumé : *On étudie, en analyse en composantes principales, l'influence sur les valeurs propres du poids d'une des unités statistiques du tableau des données. On montre comment la connaissance des coordonnées de cette unité statistique dans le repère des composantes principales permet d'améliorer les bornes à la variation des valeurs propres obtenues en [3] par l'auteur. L'aspect géométrique du problème est abordé. Les améliorations apportées sont illustrées à partir de l'exemple des poissons d'Amiard.*

Abstract : *This paper is concerned with the influence of an observation weight upon the eigenvalues in principal component analysis. Using the coordinates of the observation in the principal component coordinate system an improvment is given for the bounds to the variation of the eigenvalues obtained in [3] by the author. Geometrical aspect is pointed out. A practical illustration of the improvment, based on Amiard fishes example, is given.*

Mots clés : *Analyse en composantes principales, influence d'une unité statistique, poids d'une unité statistique, perturbation, valeurs propres.*

Indice de classification STMA : *06-070.*

Manuscrit reçu le 20.2.1986, révisé le 26.9.1986

1 - INTRODUCTION

On suppose que l'on dispose d'un tableau de données X formé de n unités statistiques (u.s.) caractérisées par les vecteurs colonnes x_i de R^p avec pour $i = 1,\dots,n$ $x_i^t = (x_{i1},\dots, x_{ip})$. Pour chaque u.s. i considérée les coordonnées x_{ik}, $k = 1,\dots,p$ correspondent aux mesures effectuées sur les p variables étudiées. Le plan d'échantillonnage adopté pour l'analyse des données est défini par la matrice nxn $D = \underset{i}{\mathrm{diag}}\ p_i$ qui regroupe les poids p_i, $i = 1,\dots,n$ des différentes u.s. ($p_i \geqslant 0$, $\sum_{i=1}^{n} p_i = 1$).

Dans l'interprétation qu'il fait des résultats de l'analyse en composantes principales (A.C.P.) de ses données le praticien est parfois conduit à s'intéresser à l'influence d'un groupe particulier d'u.s. Dans ce contexte B. ESCOFIER [5] s'est intéressée aux conséquences de la suppression ou de l'ajout d'un tel groupe. Plus récemment l'auteur a étudié en [3] l'influence d'une seule u.s. au travers d'une variation du poids qui lui est attribué et les résultats proposés tendent à répondre aux questions d'influence sur les inerties des axes avant même d'effectuer une quelconque A.C.P. des données. En pratique cependant ce sont bien souvent les résultats d'une première A.C.P. qui conduisent le statisticien à s'intéresser au rôle qu'a pû y jouer une u.s. particulière dont il connaît alors les coordonnées dans le repère des composantes principales. Nous nous proposons ici de montrer comment cette connaissance permet d'améliorer les résultats obtenus en [3].

Le paragraphe 2 est consacré à la présentation des résultats algébriques sur lesquels se fonde notre approche. Les paragraphes 3 et 4 abordent le problème de la modification du poids d'une u.s. lorsqu'on travaille respectivement sur matrice de covariance et sur matrice de corrélation. Dans le paragraphe 5 on souligne un des aspects géométriques de l'utilisation des composantes principales tandis que le dernier paragraphe illustre l'intérêt des résultats obtenus à partir de l'exemple bien connu des poissons d'Amiard dont on trouvera une présentation détaillée en [6].

2 - RESULTATS ALGEBRIQUES

Tout au long de notre exposé les valeurs propres d'une matrice carrée M d'ordre m seront notées $\lambda_\ell(M)$ $\ell = 1,\dots,m$ avec $\lambda_1(M) \geqslant \lambda_2(M) \geqslant \dots \geqslant \lambda_m(M)$.

La proposition 2.1 qui suit est un résultat classique de la théorie des perturbations que nous avons déjà utilisé en [3].

Proposition 2.1 : Soient M et N deux matrices symétriques d'ordre m. Alors les inégalités suivantes se trouvent vérifiées pour tous entiers j, k, ℓ éléments de $\{1,\dots,m\}$ vérifiant $j+k \leqslant \ell+1$.

I - Inégalités de Weyl

(1) $\lambda_\ell(M+N) \leqslant \lambda_j(M) + \lambda_k(N)$

(2) $\lambda_{m-\ell+1}(M+N) \geqslant \lambda_{m-j+1}(M) + \lambda_{m-k+1}(N)$

(3) $\lambda_m(N) + \lambda_\ell(M) \leqslant \lambda_\ell(M+N) \leqslant \lambda_\ell(M) + \lambda_1(N)$

II - Si M et N sont semi définies positives

(4) $\lambda_\ell(MN) \leqslant \lambda_j(M)\, \lambda_k(N)$

(5) $\lambda_{m-\ell+1}(MN) \geqslant \lambda_{m-j+1}(M)\, \lambda_{m-k+1}(N)$

(6) $\lambda_m(N)\, \lambda_\ell(M) \leqslant \lambda_\ell(MN) \leqslant \lambda_\ell(M)\, \lambda_1(N)$

Remarque : Pour plus de détails concernant les inégalités de Weyl on pourra se référer à [7]. Les inégalités II ont été obtenues en [1] lorsque N est supposée définie positive. Le résultat s'étend facilement au cas où N est seulement semi définie positive (voir [2]). On notera enfin que les inégalités (3) et (6) que nous employerons le plus souvent par la suite s'obtiennent comme cas particulier des inégalités (1), (2) et (4), (5) respectivement.

La proposition 2.4 qui suit avec son corollaire constitue un résultat technique spécifique du problème abordé ici. Ce résultat repose lui même sur le lemme 2.2 dont on pourra trouver la démonstration en [4], ainsi que sur le lemme 2.3 qui est immédiat. Dans les énoncés la notation $\langle .,. \rangle$ désigne le produit scalaire habituel sur R^m et $||.||$ la norme associée.

Lemme 2.2 : Soient u et v deux vecteurs colonnes de R^m linéairement indépendants, q, r, s trois réels vérifiant $qs-r^2 \neq 0$. Alors la matrice $M = q\, u\, u^t + r(uv^t + vu^t) + s\, v\, v^t$ est de rang deux et ses valeurs propres non nulles s'expriment sous la forme $\frac{1}{2}\left[\alpha \pm \sqrt{\alpha^2 - 4(qs-r^2)\,\beta}\right]$ avec $\alpha = q\, ||u||^2 + 2r \langle u,v \rangle + s\, ||v||^2$ et $\beta = ||u||^2\, ||v||^2 - \langle u,v \rangle^2$. Elles sont de signes opposés si $qs-r^2 < 0$ et de même signe si $qs-r^2 > 0$.

Lemme 2.3 : Soit u un vecteur colonne de R^m et q un réel. Alors la matrice $M = q\, u\, u^t$ est de rang un avec pour valeur propre non nulle $q\, ||u||^2$.

Proposition 2.4 : Soit $A = \underset{\ell}{\text{diag}}\ \alpha_\ell$ la matrice d'ordre m d'un opérateur dans la base naturelle $(e_1, \ldots, e_m)$ de R^m. On suppose $\alpha_1 > \alpha_2 > \ldots > \alpha_m$. Soit (I,J) une partition de l'ensemble $E = \{1,\ldots,m\}$ en deux sous ensembles non vides, u un vecteur de R^m défini par $u = \sum_{k=1}^{m} u_k\, e_k$, u_I et u_J les vecteurs définis par $u_I = \sum_{i \in I} u_i\, e_i$, $u_J = \sum_{j \in J} u_j\, e_j$. Alors les valeurs propres de la matrice $B = A + u\, u^t$ vérifient pour tout $i \in I$:

$$\lambda_i(B) \leqslant \alpha_i + \frac{1}{2}\left[||u_I||^2 + \sqrt{||u_I||^4 + 4||u_I||^2\, ||u_J||^2}\right]$$

si on a $\alpha_i - \alpha_{i+1} > ||u_J||^2$ (pour $i \neq m$ lorsque $m \in I$).

Preuve : u se décompose sous la forme $u = u_I + u_J$. On a alors

$$B = A + u\, u^t = (A + u_J\, u_J^t) + (u_I\, u_I^t + u_I\, u_J^t + u_J\, u_I^t)$$

A partir de cette égalité la proposition 2.1.I permet d'écrire pour tout $\ell \in \{1,\ldots,m\}$

$$\lambda_\ell(B) \leqslant \lambda_\ell(A+u_Ju_J^t) + \lambda_1(u_Iu_I^t + u_Iu_J^t + u_Ju_I^t) \qquad (1)$$

Comme par construction on a $\langle u_I, u_J\rangle = 0$ on obtient à partir du lemme 2.2 :

$$\lambda_1(u_Iu_I^t + u_Iu_J^t + u_Ju_I^t) = \frac{1}{2}\left[||u_I||^2 + \sqrt{||u_I||^4 + 4||u_I||^2\ ||u_J||^2}\right] \quad (2)$$

D'après les relations (1) et (2) il suffit alors de prouver que lorsque $i \in I$ on a $\lambda_i(A + u_J\, u_J^t) = \lambda_i(A) = \alpha_i$ pour obtenir le résultat désiré.

Il est facile de voir que α_i est valeur propre de $A + u_Ju_J^t$ lorsque $i \in I$ puisqu'on a alors $\langle u_J, e_i\rangle = 0$ d'où l'on déduit :

$$(A + u_Ju_J^t)\, e_i = A\, e_i + \langle u_J, e_i\rangle\, u_J = A\, e_i = \alpha_i\, e_i$$

Pour montrer que α_i est bien la <u>i-ième</u> valeur propre de $A + u_Ju_J^t$ on utilise le fait que les α_i ont été supposées distinctes et la condition $\alpha_i - \alpha_{i+1} > ||u_J||^2$. En appliquant la proposition 2.1 I à $A + u_Ju_J^t$ on obtient alors, compte tenu du lemme 2.3, pour $i \neq 1$ et $i \neq m$:

$$\lambda_{i+1}(A + u_Ju_J^t) \leqslant \alpha_{i+1} + ||u_J||^2 < \alpha_i < \alpha_{i-1} \leqslant \lambda_{i-1}(A + u_Ju_J^t)$$

$$\text{i.e.}\quad \lambda_{i+1}(A + u_Ju_J^t) < \alpha_i < \lambda_{i-1}(A + u_Ju_J^t)$$

pour $i = 1$: $\alpha_1 > \lambda_2(A+u_Ju_J^t)$ et pour $i = m$: $\alpha_m < \lambda_{m-1}(A+u_Ju_J^t)$

Ces inégalités montrent que dans tous les cas α_i est bien la i-ième valeur propre de $A + u_Ju_J^t$.

Corollaire 2.5 : Pour toutes matrices $A = \operatorname{diag}_{\ell}\, \alpha_\ell$ avec $\alpha_1 > \ldots > \alpha_m$ et $B = A + u\, u^t$ avec $u^t = (u_1, \ldots, u_m)$ on a l'inégalité :

$$\lambda_\ell(B) \leqslant \alpha_\ell + \frac{1}{2}\left[u_\ell^2 + \sqrt{u_\ell^4 + 4\, u_\ell^2\,(||u||^2 - u_\ell^2)}\right] \qquad (1)$$

pour tout $\ell \in \{1,\ldots,m\}$ vérifiant $\alpha_\ell - \alpha_{\ell+1} > ||u||^2 - u_\ell^2$ lorsque $\ell \neq m$

et l'inégalité :

$$\lambda_\ell(B) \leqslant \alpha_\ell + \frac{1}{2}\left[(u_\ell^2+u_k^2) + \sqrt{(u_\ell^2+u_k^2) + 4(u_\ell^2+u_k^2)(||u||^2 - (u_\ell^2+u_k^2))}\right] \qquad (2)$$

pour tous ℓ et k éléments distincts de $\{1,...,m\}$ vérifiant
$\alpha_\ell - \alpha_{\ell+1} > ||u||^2 - (u_\ell^2 + u_k^2)$ lorsque $\ell \neq m$.

Preuve : On applique la proposition 2.4 avec pour (1) :

$$I = \{\ell\} \quad \text{et} \quad J = \{1,...,m\} - I$$

et pour (2) : $\quad I = \{\ell,k\}$ et $J = \{1,...,m\} - I$

Remarques : 1 - Les résultats de la proposition 2.5 et de son corollaire se transposent sans difficulté au cas de $B = A - u\,u^t$. En ce qui concerne la proposition on obtient alors pour $i \in I$:

$$\lambda_i(B) \geqslant \alpha_i - \frac{1}{2}\left[||u_I||^2 + \sqrt{||u_I||^4 + 4\,||u_I||^2\,||u_J||^2}\right]$$

si l'on a $\alpha_{i-1} - \alpha_i > ||u_J||^2$ (pour $i \neq 1$ lorsque $1 \in I$). Les inégalités (1) et (2) du corollaire deviennent respectivement pour leur part :

$$\lambda_\ell(B) \geqslant \alpha_\ell - \frac{1}{2}\left[u_\ell^2 + \sqrt{u_\ell^4 + 4\,u_\ell^2\,(||u||^2 - u_\ell^2)}\right]$$

si l'on a $\quad \alpha_{\ell-1} - \alpha_\ell > ||u||^2 - u_\ell^2$ lorsque $\ell \neq 1$ et :

$$\lambda_\ell(B) \geqslant \alpha_\ell - \frac{1}{2}\left[(u_\ell^2+u_k^2) + \sqrt{(u_\ell^2+u_k^2)^2 + 4(u_\ell^2+u_k^2)(||u||^2 - (u_\ell^2+u_k^2))}\right]$$

si l'on a $\quad \alpha_{\ell-1} - \alpha_\ell > ||u||^2 - (u_\ell^2+u_k^2)$ lorsque $\ell \neq 1$.

2 - Une application directe de la proposition 2.1 I permet d'obtenir, dans le cas $B = A + u\,u^t$: $\lambda_\ell(B) \leqslant \lambda_\ell(A) + ||u||^2$, dans le cas $B = A - u\,u^t$: $\lambda_\ell(B) \geqslant \lambda_\ell(A) - ||u||^2$. La proposition 2.6 montre l'amélioration apportée par les résultats qui précèdent :

Proposition 2.6 : Avec les notations de la proposition 2.4 on a :

$$\frac{1}{2}\left[||u_I||^2 + \sqrt{||u_I||^4 + 4\,||u_I||^2\,||u_J||^2}\right] \leqslant ||u||^2$$

Preuve : on a :

$$\frac{1}{2}\left[\|u_I\|^2+\sqrt{\|u_I\|^4+4\|u_I\|^2\,\|u_J\|^2}\right] = \frac{1}{2}\left[\|u_I\|^2+\sqrt{\|u_I\|^4+4\|u_I\|^2(\|u\|^2-\|u_I\|^2)}\right]$$

Considérons alors la fonction f définie par :

$$f : \left[0, \, ||u||^2\right] \longrightarrow R^+$$
$$x \longmapsto \frac{1}{2}\left[x + \sqrt{x^2+4x(||u||^2 - x)}\right]$$

Il est facile de voir que f est croissante concave sur son intervalle de définition. Alors comme $||u_I||^2 \leqslant ||u||^2$ on en déduit $f(||u_I||^2) \leqslant f(||u||^2)$ ce qui correspond au résultat proposé.

Remarques : 1 - Il est immédiat de voir que l'application g définie sur $\left[0, \, ||u||^2\right]$ par $g(x) = f(x)/x$ est décroissante. La concavité de f et la décroissance de g permettent de mieux cerner la qualité de l'amélioration qu'apporte le terme $\frac{1}{2}\left[||u_I||^2 + \sqrt{||u_I||^4 + 4\,||u_I||^2\,||u_J||^2}\right]$ par rapport à $||u||^2$. En fait la concavité de f contrarie une amélioration rapide lorsque $||u_I||^2$ décroit à partir de $||u||^2$.

2 - La croissance de f permet de voir qu'il est toujours préférable d'utiliser la relation (1) du corollaire 2.5 plutôt que la relation (2). En fait on n'utilise la relation (2) que lorsque la condition $\alpha_\ell - \alpha_{\ell+1} > ||u||^2 - u_\ell^2$ n'est pas vérifiée. On choisit alors le second indice k de manière à avoir $\alpha_\ell - \alpha_{\ell+1} > ||u||^2 - (u_\ell^2 + u_k^2)$ tout en ayant u_k^2 le plus petit possible.

3 - INFLUENCE DU POIDS D'UNE UNITE STATISTIQUE SUR LES VALEURS PROPRES DE LA MATRICE DE COVARIANCE

On aborde à présent le problème qui nous intéresse directement, à savoir celui des conséquences sur les valeurs propres d'une modification du poids p_j de l'u.s. j en un poids $\tilde{p}_j$ quelconque $(0 \leqslant \tilde{p}_j < 1)$. Les poids p_i $i = 1,\dots,n$ $i \neq j$ des autres u.s. sont ici seulement normalisés en $\tilde{p}_i = \alpha_j\, p_i$ avec $\alpha_j = (1-\tilde{p}_j)/(1-p_j)$ afin d'avoir $\sum_{i=1}^{n} \tilde{p}_i = 1$. Dans ce paragraphe on se place tout d'abord dans le cadre de l'A.C.P. du triplet (X, Q, D) avec $Q = I_p$ métrique identité choisie dans R^p pour mesurer les distances entre u.s., ce qui conduit à s'intéresser aux valeurs propres de la matrice de covariance.

3.1. On sait (voir [3]) que suite à la modification de poids la transformation de la matrice de covariance initiale V en $\tilde{V}$ est définie par la relation :

$$\tilde{V} = \alpha_j \left[V + \beta_j \, (x_j - g)(x_j - g)^t \right] \tag{3.1}$$

avec $\beta_j = 1 - \alpha_j$, g désignant le centre de gravité des u.s. pour les poids initiaux ($g = \sum_{i=1}^{n} p_i \, x_i$).

On obtient alors en particulier par application de la proposition 2.1.I les encadrements suivants pour les valeurs propres de $\tilde{V}$.

- dans le cas d'une augmentation de poids ($\tilde{p}_j > p_j$) i.e. $\beta_j > 0$, pour $\ell = 1,\dots,p$

$$\alpha_j \, \lambda_\ell(V) \leqslant \lambda_\ell(\tilde{V}) \leqslant \alpha_j \left\{ \lambda_\ell(V) + \beta_j \, ||x_j - g||^2 \right\} \tag{3.2}$$

- dans le cas d'une diminution de poids ($\tilde{p}_j < p_j$) i.e. $\beta_j < 0$, pour $\ell = 1,\dots,p$

$$\alpha_j \left\{ \lambda_\ell(V) + \beta_j \, ||x_j - g||^2 \right\} \leqslant \lambda_\ell(\tilde{V}) \leqslant \alpha_j \, \lambda_\ell(V) \tag{3.3}$$

3.2. Lorsque les valeurs propres de V sont supposées distinctes, nous allons voir que l'utilisation des composantes principales permet d'améliorer la précision de la borne $\alpha_j \left\{ \lambda_\ell(V) + \beta_j \, ||x_j - g||^2 \right\}$ intervenant dans les expressions (3.2) et (3.3).

En effet désignons par F la matrice pxp dont les colonnes (facteurs principaux de l'A.C.P.) sont les vecteurs propres de V normés à l'unité ($F^tF = I$).

Avec $\Lambda = \underset{\ell}{\mathrm{diag}}(\lambda_\ell(V))$ on a : $V\,F = F\,\Lambda$

Introduisons enfin le vecteur $z_j^t = (z_{j1}, \dots, z_{jp}) = (x_j - g)^t \, F$ des coordonnées de l'u.s. j dans le repère des composantes principales de l'A.C.P. pour les poids initiaux.

En multipliant (3.1) à gauche par F^t et à droite par F on constate que $\tilde{V}$ a les mêmes valeurs propres que :

$$\alpha_j \left[\Lambda + \beta_j z_j z_j^t\right] \tag{3.4}$$

La matrice Λ étant diagonale et les valeurs propres $\lambda_\ell(V)$ supposées distinctes on obtient par une application directe de l'inégalité (1) du corollaire 2.5 le résultat suivant :

Si $\beta_j > 0$, pour tout $\ell \in \{1,\dots,p\}$ vérifiant $\lambda_\ell(V) - \lambda_{\ell+1}(V) > \beta_j \left[||z_j||^2 - z_{j\ell}^2\right]$ lorsque $\ell \neq p$:

$$\lambda_\ell(\tilde{V}) \leq \alpha_j \left\{\lambda_\ell(V) + \frac{1}{2}\beta_j \left[z_{j\ell}^2 + \sqrt{z_{j\ell}^4 + 4\, z_{j\ell}^2\,(||z_j||^2 - z_{j\ell}^2)}\right]\right\} \tag{3.5}$$

En référence à la remarque qui suit le corollaire 2.5, on obtient également à partir de (3.4) :

- Si $\beta_j < 0$, pour tout $\ell \in \{1,\dots,p\}$ vérifiant $\lambda_{\ell-1}(V) - \lambda_\ell(V) > -\beta_j \left[||z_j||^2 - z_{j\ell}^2\right]$ lorsque $\ell \neq 1$:

$$\lambda_\ell(\tilde{V}) \geq \alpha_j \left\{\lambda_\ell(V) + \frac{1}{2}\beta_j \left[z_{j\ell}^2 + \sqrt{z_{j\ell}^4 + 4\, z_{j\ell}^2\,(||z_j||^2 - z_{j\ell}^2)}\right]\right\} \tag{3.6}$$

Remarques : 1 - En utilisant la proposition 2.5 avec $I = \{\ell\}$ et $J = \{1,\dots,p\} - I$, on obtient :

$$\frac{1}{2}\left[z_{j\ell}^2 + \sqrt{z_{j\ell}^2 + 4\, z_{j\ell}^2\,(||z_j||^2 - z_{j\ell}^2)}\right] \leq ||z_j||^2 = ||x_j - g||^2$$

L'inégalité (3.5) (resp. (3.6)) fournit donc une borne plus précise que la borne supérieure (resp. inférieure) obtenue en (3.2) (resp. en (3.3)).

2 - Lorsque $z_{j\ell} = 0$, les inégalités (3.5) et (3.6) prennent respectivement la forme suivante :

$$\text{si } \beta_j > 0 \qquad \lambda_\ell(\tilde{V}) \leq \alpha_j\, \lambda_\ell(V) \tag{3.7}$$

$$\text{si } \beta_j < 0 \qquad \lambda_\ell(\tilde{V}) \geq \alpha_j\, \lambda_\ell(V) \tag{3.8}$$

La comparaison de (3.7) et du membre de gauche de (3.2) dans le cas $\beta_j > 0$, celle de (3.3) et du membre de droite de (3.3) pour $\beta_j < 0$ permettent de voir que l'on a alors $\lambda_\ell(\tilde{V}) = \alpha_j \lambda_\ell(V)$.

3 - On notera que l'utilisation de la relation (2) du corollaire 2.5 conduit aussi à une majoration de $\lambda_\ell(\tilde{V})$ dans le cas $\beta_j > 0$ et à une minoration dans le cas $\beta_j < 0$. Pour ne pas alourdir la présentation nous ne détaillons pas ce type de formulation qui fait intervenir une deuxième coordonnée z_{jk} en plus de $z_{j\ell}$. Les bornes obtenues sont moins précises que celles de (3.5) et (3.6) (remarque 2 à la suite de la proposition 2.6) mais peuvent présenter un intérêt pratique qui sera illustré au dernier paragraphe.

4 - Dans l'interprétation qu'ils font des résultats les praticiens ont l'habitude d'utiliser les contributions des observations à l'inertie des axes. Dans cette optique il est intéressant de noter qu'il est facile de reexprimer la borne $\alpha_j \left\{\lambda_\ell(V) + \frac{1}{2}\beta_j \left[z_{j\ell}^2 + \sqrt{z_{j\ell}^4 + 4\, z_{j\ell}^2\,(||z_j||^2 - z_{j\ell}^2)}\right]\right\}$ intervenant dans les relations (3.5) et (3.6) en fonction de la contribution absolue $c_{j\ell} = p_j(z_{j\ell}^2)/\lambda_\ell(V)$ de l'u.s. j à l'axe ℓ. On obtient alors l'expression :

$$\alpha_j\, \lambda_\ell(V) \left\{1 + \frac{1}{2}\,(\beta_j/p_j)\, \left[c_{j\ell} + \sqrt{c_{j\ell}^2 + 4c_{j\ell}(p_j(\lambda_\ell(V))^{-1}||z_j||^2 - c_{j\ell})}\right]\right\}$$

On note en particulier que $\left[c_{j\ell} + \sqrt{c_{j\ell}^2 + 4c_{j\ell}(p_j(\lambda_\ell(V))^{-1}||z_j||^2 - c_{j\ell})}\right]$ est une fonction croissante par rapport à $c_{j\ell}$. On confirme ainsi l'intuition naturelle selon laquelle la variation de la valeur propre de rang ℓ doit a priori pouvoir être d'autant plus importante que l'u.s. j a initialement une contribution importante à l'axe ℓ.

Remarquons enfin que dans le cas particulier (important en pratique) de la suppression d'une u.s. j avec initialement tous les poids égaux à 1/n on a $\alpha_j = \beta_j/p_j = \frac{n}{n-1}$ ce qui conduit pour la borne étudiée à l'expression :

$$(n/(n-1))\lambda_\ell \left\{1+(1/2)(n/(n-1))\left[c_{j\ell} + \sqrt{c_{j\ell}^2+4c_{j\ell}(((\lambda_\ell(V))^{-1}||z_j||^2/n)-c_{j\ell})}\right]\right\}$$

4 - INFLUENCE DU POIDS D'UNE UNITE STATISTIQUE SUR LES VALEURS PROPRES DE LA MATRICE DE CORRELATION

On aborde ici le problème du paragraphe précédent dans le cadre de l'A.C.P. du triplet (X, Q, D) avec Q métrique diagonale dans R^p dont les éléments diagonaux sont les inverses des variances des variables, ce qui conduit à s'intéresser aux valeurs propres de la matrice de corrélation.

4.1. Désignons par $y_j^t = (y_{j1}, \ldots, y_{jp})$ le vecteur des coordonnées de l'u.s. j centrée réduite pour les poids initiaux. La transformation du poids de cette u.s. entraine (voir [3]) une modification de la matrice de corrélation initiale R en $\tilde{R}$ définie par l'expression :

$$\tilde{R} = T\,(R + \beta_j\, y_j\, y_j^t)\, T \tag{4.1}$$

où $T = \operatorname{diag}_k t_k$ avec $t_k = (1 + \beta_j\, y_{jk}^2)^{-1/2}$. On remarque d'après (4.1) que $\tilde{R}$ a les mêmes valeurs propres que $(R + \beta_j\, y_j\, y_j^t)\, T^2$ et l'on en déduit les encadrements suivants en utilisant la partie II de la proposition 2.1.

$$\left\{\min_k(1+\beta_j y_{jk}^2)^{-1}\right\}\left\{\lambda_\ell(R+\beta_j y_j y_j^t)\right\} \leqslant \lambda_\ell(\tilde{R}) \leqslant \left\{\lambda_\ell(R+\beta_j y_j y_j^t)\right\}\left\{\max_k(1+\beta_j y_{jk}^2)^{-1}\right\} \tag{4.2}$$

En utilisant ensuite la partie I de la même proposition pour $R + \beta_j y_j y_j^t$, on obtient, compte tenu du lemme 2.3, pour tout $\ell \in \{1,\ldots,p\}$:

- si $\beta_j > 0$

$$\left\{\min_k(1+\beta_j y_{jk}^2)^{-1}\right\}\left\{\lambda_\ell(R)\right\} \leqslant \lambda_\ell(\tilde{R}) \leqslant \left\{\lambda_\ell(R)+\beta_j\,\|y_j\|^2\right\}\left\{\max_k(1+\beta_j y_{jk}^2)^{-1}\right\} \tag{4.3}$$

- si $\beta_j < 0$

$$\left\{\min_k(1+\beta_j y_{jk}^2)^{-1}\right\}\left\{\lambda_\ell(R)+\beta_j\,\|y_j\|^2\right\} \leqslant \lambda_\ell(\tilde{R}) \leqslant \left\{\lambda_\ell(R)\right\}\left\{\max_k(1+\beta_j y_{jk}^2)^{-1}\right\} \tag{4.4}$$

4.2. Ces inégalités (4.3) et (4.4) peuvent cependant être améliorées en procédant avec l'expression $R + \beta_j y_j y_j^t$ comme il a été fait avec l'expression (3.1) $\tilde{V} = \alpha_j\left[V+\beta_j(x_j-g)(x_j-g)^t\right]$, dans le cas de la matrice de covariance, au paragraphe précédent.

En désignant par $z_j^t = (z_{j1}, \ldots, z_{jp})$ le vecteur des coordonnées de l'u.s. j dans le repère des composantes principales de l'A.C.P. des données centrées réduites pour les poids initiaux , on obtient alors en repartant de (4.2) :

- dans le cas $\beta_j > 0$:

pour tout $\ell \in \{1,\ldots,p\}$ vérifiant $\lambda_\ell(R)-\lambda_{\ell+1}(R) > \beta_j\left[||z_j||^2 - z_{j\ell}^2\right]$ lorsque $\ell \neq p$:

$$\lambda_\ell(\tilde{R}) \leqslant \left\{\lambda_\ell(R) + \frac{1}{2}\beta_j\left[z_{j\ell}^2 + \sqrt{z_{j\ell}^4 + 4\,z_{j\ell}^2(||z_j||^2 - z_{j\ell}^2)}\right]\right\}\left\{\max_k(1+\beta_j y_{jk}^2)^{-1}\right\} \tag{4.5}$$

- dans le cas $\beta_j < 0$:

pour tout $\ell \in \{1,\ldots,p\}$ vérifiant $\lambda_{\ell-1}(R)-\lambda_\ell(R) > -\beta_j\left[||z_j||^2 - z_{j\ell}^2\right]$ lorsque $\ell \neq 1$:

$$\lambda_\ell(\tilde{R}) \geqslant \left\{\lambda_\ell(R) + \frac{1}{2}\beta_j\left[z_{j\ell}^2 + \sqrt{z_{j\ell}^4 + 4\,z_{j\ell}^2(||z_j||^2 - z_{j\ell}^2)}\right]\right\}\left\{\min_k(1+\beta_j y_{jk}^2)^{-1}\right\} \tag{4.6}$$

Les remarques 1 à 4 faites dans le cadre de l'étude de l'A.C.P. sur matrice de covariance se transposent sans difficulté au cadre de la matrice de corrélation. En particulier l'amélioration qu'apportent les inégalités (4.5) et (4.6) par rapport aux encadrements (4.2) et (4.3) tient toujours au résultat de la proposition 2.6 selon lequel :

$$\frac{1}{2}\left[z_{j\ell}^2 + \sqrt{z_{j\ell}^4 + 4\,z_{j\ell}^2\,(||z_j||^2 - z_{j\ell}^2)}\right] \leqslant ||z_j||^2 = ||y_j||^2$$

5 - ASPECT GEOMETRIQUE

Etant donné un réel positif θ, on donne en [3] des conditions géométriques sur l'u.s. dont on perturbe le poids, qui assurent que cette modification n'entrainera pas de perturbation des valeurs propres d'amplitude supérieure à θ. L'utilisation des composantes principales permet de compléter ces conditions géométriques par les résultats que nous présentons ici. Cette présentation est faite dans le cadre de l'A.C.P. sur matrice de corrélation qui est la plus usitée en pratique mais la transposition des résultats au cadre de l'A.C.P. sur matrice de covariance se fait de manière immédiate.

Comme en [3] on utilise le lemme suivant :

Lemme 4.1 :

(i) si $\beta_j > 0$ $(\tilde{p}_j > p_j)$ $\qquad \lambda_1(T^2) = \max\limits_k \left\{(1 + \beta_j\, y_{jk}^2)^{-1}\right\} \leqslant 1$

(ii) si $\beta_j < 0$ $(\tilde{p}_j < p_j)$ $\qquad \lambda_p(T^2) = \min\limits_k \left\{(1 + \beta_j\, y_{jk}^2)^{-1}\right\} \geqslant 1$

On étudie le cas d'une augmentation de poids $(\beta_j > 0)$. Les résultats concernant une diminution de poids pourront s'obtenir par un raisonnement similaire à celui qui va être présenté.

En utilisant la relation (i) du lemme précédent conjointement au membre de droite de la relation (4.2) on obtient pour tout $\ell \in \{1,\dots,p\}$

$$\lambda_\ell(\tilde{R}) \leqslant \lambda_\ell(R + \beta_j\, y_j\, y_j^t) \tag{5.1}$$

En posant $\Lambda = \operatorname{diag}_\ell(\lambda_\ell(R))$, il est facile de voir que la matrice $R + \beta_j\, y_j\, y_j^t$ a mêmes valeurs propres que $\Lambda + \beta_j\, z_j\, z_j^t$ (avec z_j défini comme au paragraphe précédent) et l'inégalité (5.1) peut donc s'écrire :

$$\lambda_\ell(\tilde{R}) \leqslant \lambda_\ell(\Lambda + \beta_j\, z_j\, z_j^t) \tag{5.2}$$

Or, d'après la relation

$$\Lambda + \beta_j\, z_j\, z_j^t = \Lambda^{1/2}(I + \beta_j\, \Lambda^{-1/2}\, z_j\, z_j^t\, \Lambda^{-1/2})\, \Lambda^{1/2}$$

(avec $\Lambda^{1/2} = \operatorname{diag}(\sqrt{\lambda_\ell(R)})$ et $\Lambda^{-1/2} = \operatorname{diag}(1/\sqrt{\lambda_\ell(R)})$ on voit que $\Lambda + \beta_j\, z_j\, z_j^t$ a mêmes valeurs propres que $\Lambda\,(I + \beta_j\, \Lambda^{-1/2}\, z_j\, z_j^t\, \Lambda^{-1/2})$.
On remarque facilement en ce qui concerne la matrice $I + \beta_j\, \Lambda^{-1/2}\, z_j\, z_j^t\, \Lambda^{-1/2}$ que la plus grande valeur propre est égale à $1 + \beta_j\, z_j^t\, \Lambda^{-1}\, z_j$ $(\Lambda^{-1/2}\, z_j$ est un vecteur propre associé ; les autres valeurs propres sont toutes égales à l'unité, tout élément de l'hyperplan orthogonal à $\Lambda^{-1/2}\, z_j$ pour le produit scalaire habituel étant un vecteur propre associé). On obtient alors en utilisant la proposition 2.1 :

$$\lambda_\ell(\Lambda + \beta_j\, z_j\, z_j^t) \leqslant \left\{\lambda_\ell(R)\right\} \left\{1 + \beta_j\, z_j^t\, \Lambda^{-1}\, z_j\right\}$$

soit compte tenu de (5.2) :

$$\lambda_\ell(\tilde{R}) \leqslant \left\{\lambda_\ell(R)\right\} \left\{1 + \beta_j \, z_j^t \, \Lambda^{-1} \, z_j\right\} \tag{5.3}$$

expression qui peut encore s'écrire sous la forme :

$$\lambda_\ell(\tilde{R}) \leqslant \left\{\lambda_\ell(R)\right\} \left\{1 + \beta_j \sum_{k=1}^{p} \frac{z_{jk}^2}{\lambda_k(R)}\right\} \tag{5.4}$$

On peut alors énoncer le résultat suivant :

Proposition 5.2 : Soient $\ell \in \{1,\ldots,p\}$ et θ un réel positif donné.

Lorsqu'on augmente le poids d'une u.s. j de la valeur p_j à $\tilde{p}_j$ ($\beta_j > 0$), il suffit que cette u.s. appartienne à l'intérieur de l'ellipsoïde de R^p défini dans le repère des composantes principales par

$$\sum_{k=1}^{p} \frac{x_k^2}{a_k^2} = 1 \qquad \text{avec} \qquad a_k = \left\{\frac{\theta}{\beta_j} \times \frac{\lambda_k(R)}{\lambda_\ell(R)}\right\}^{1/2} \qquad \forall k = 1,\ldots,p$$

pour que l'augmentation éventuelle de la ℓ-ième valeur propre reste d'amplitude inférieure à θ i.e. pour que l'on ait :

$$\lambda_\ell(\tilde{R}) \leqslant \lambda_\ell(R) + \theta$$

Preuve : Si l'u.s. j appartient à l'ellipsoïde en question, le vecteur $z_j^t = (z_{j1}, \ldots, z_{jp})$ de ses coordonnées dans le repère des composantes principales vérifie :

$$\begin{aligned}
\sum_{k=1}^{p} \frac{z_{jk}^2}{a_k^2} \leqslant 1 \iff & \frac{\beta_j}{\theta} \lambda_\ell(R) \sum_{k=1}^{p} \frac{z_{jk}^2}{\lambda_k(R)} \leqslant 1 \\
\iff & \lambda_\ell(R) + \beta_j \, \lambda_\ell(R) \sum_{k=1}^{p} \frac{z_{jk}^2}{\lambda_k(R)} \leqslant \lambda_\ell(R) + \theta \\
\iff & \lambda_\ell(R) \left[1 + \beta_j \sum_{k=1}^{p} \frac{z_{jk}^2}{\lambda_k(R)}\right] \leqslant \lambda_\ell(R) + \theta \\
\iff & \lambda_\ell(\tilde{R}) \leqslant \lambda_\ell(R) + \theta \quad \text{(en utilisant (5.4))}
\end{aligned}$$

Remarques : 1 - Ce résultat est à mettre en parallèle avec celui obtenu en [3] selon lequel l'augmentation éventuelle de chaque valeur propre ne peut dépasser θ si l'u.s. centrée réduite initiale y_j se trouve dans la sphère $S_{\theta j}$ de R^p de rayon $r_{\theta j} = \sqrt{\theta/\beta_j}$ centrée sur l'origine.

On remarquera que l'ellipsoïde relatif à la première valeur propre est contenu dans la sphère $S_{\theta j}$ ($a_k \leqslant \sqrt{\theta/\beta_j}$ $\forall k = 1,\ldots,p$ car $|\lambda_k(R)/\lambda_1(R)| \leqslant 1$) alors que cette dernière est elle-même contenue dans l'ellipsoïde correspondant à la dernière valeur propre ($a_k \geqslant \sqrt{\theta/\beta_j}$ $k = 1,\ldots,p$). On notera d'autre part que tous les ellipsoïdes sont homothétiques.

2 - On obtient des résultats analogues en ce qui concerne l'amplitude d'une éventuelle diminution des valeurs propres dans l'étude d'une diminution du poids de l'u.s. j ($\beta_j < 0$).

6 - ILLUSTRATION DES RESULTATS

Nous avons choisi pour illustrer les résultats présentés l'exemple des poissons d'Amiard que l'ouvrage de PAGES et CAILLIEZ [6] a rendu classique en analyse des données. Cet exemple nous a déjà servi en [3]. Rappelons simplement qu'il s'agit d'étudier un groupe de 24 poissons placés dans une eau radioactive et pour lesquels on mesure des caractéristiques de taille et de radioactivité de différents organes. Le poisson 17 mort au cours de l'expérimentation n'est pas pris en compte dans l'étude statistique. Les 16 variables étudiées étant non homogènes quant à leur nature et leur variance on travaille sur la matrice de corrélation.

Notre objectif est ici de mettre en évidence sur un exemple l'amélioration qu'apporte les inégalités du type (4.5), (4.6) par rapport à (4.3), (4.4). Pour ce faire nous présentons de manière comparative le comportement des bornes de (4.4) et (4.6) lorsqu'on envisage successivement la suppression de chacune des u.s. (ici les 23 poissons participant à l'étude).

Pour ne pas alourdir notre exposé nous limitons notre présentation au comportement des deux premières valeurs propres qui initialement prennent les valeurs suivantes

$$\lambda_1(R) = 7.607 \qquad \lambda_2(R) = 3.763$$

lorsque des poids tous égaux à 1/23 sont attribués aux u.s. La suppression d'une u.s. j quelconque ($p_j = 1/23$, $\tilde{p}_j = 0$) conduit à attribuer pour $i \neq j$ un nouveau poids $\tilde{p}_i = 1/22$ aux autres u.s. On a alors :

$$\alpha_j = 23/22 = 1.0454545 \qquad \beta_j = - 1/22 = - 0.0454545$$

Les résultats sont présentés dans le tableau qui suit. Pour chaque u.s. supprimée on donne le carré de la distance de cette u.s. au centre de gravité initial ($||y_j||^2$), les bornes inférieures m_ℓ et supérieure M_ℓ ($\ell = 1,2$) obtenues par l'encadrement (4.4) pour la variation des deux plus grandes valeurs propres, les coordonnées $z_{j\ell}$ ($\ell = 1,2$) de cette u.s. sur les deux premières composantes principales initiales, les bornes inférieures b_ℓ ($\ell = 1,2$) obtenues par l'inégalité (4.6) pour la variation des deux premières valeurs propres. Pour la deuxième valeur propre l'obtention de la borne b_2 est soumise à la condition :

$$\lambda_1(R) - \lambda_2(R) > - \beta_j \left[||z_j||^2 - z_{j2}^2\right] \Longleftrightarrow 84.563 > \left[||y_j||^2 - z_{j2}^2\right]$$

($||z_j|| = ||y_j||$). Il est immédiat de constater que cette condition est toujours vérifiée dans notre exemple puisque la plus grande valeur de $||y_j||^2$ est obtenue pour l'u.s. 20 et que l'on a $||y_{20}||^2 = 49.921$. Il est donc dans tous les cas légitime d'utiliser l'expression (4.6) qui donne la borne b_2.

Il convient pour apprécier l'amélioration qu'apporte la relation (4.6) par rapport à la relation (4.4) de comparer pour chaque u.s. supprimée b_1 à m_1 et b_2 à m_2. On a toujours naturellement $b_1 \geqslant m_1$ et $b_2 \geqslant m_2$ d'après la proposition 2.6.

u.s. j	$\|y_j\|^2$	m_1	M_1	m_2	M_2	z_{j1}	z_{j2}	b_1	b_2
1	32.563	6.233	10.266	2.323	5.079	- 5.207	- 1.642	6.270	3.346
2	18.451	6.772	8.886	2.926	4.396	- 3.974	- 0.093	6.788	3.747
3	22.368	6.591	8.956	2.747	4.431	- 4.611	- 0.411	6.593	3.671
4	27.705	6.351	9.247	2.505	4.575	- 5.087	- 0.577	6.356	3.620
5	17.309	6.821	8.990	2.977	4.447	1.371	3.657	7.316	3.011
6	11.039	7.145	8.111	3.280	4.013	0.504	3.090	7.568	3.288
7	14.342	7.068	8.653	3.162	4.281	0.388	3.646	7.659	3.165
8	15.907	6.909	9.123	3.051	4.513	1.500	2.925	7.325	3.171
9	3.538	7.446	7.910	3.603	3.913	0.345	1.192	7.575	3.646
10	5.776	7.345	8.297	3.501	4.105	0.700	- 0.654	7.522	3.684
11	5.143	7.374	8.281	3.530	4.096	- 0.525	0.667	7.548	3.687
12	5.641	7.351	8.533	3.507	4.221	1.065	0.440	7.476	3.712
13	12.051	7.059	10.862	3.216	5.374	0.166	- 0.769	7.581	3.631
14	14.364	6.954	8.532	3.110	4.221	2.680	1.608	7.079	3.447
15	2.808	7.480	7.955	3.636	3.935	- 0.967	0.437	7.522	3.727
16	6.171	7.327	8.333	3.483	4.122	- 1.284	- 0.407	7.440	3.714
18	22.975	6.566	10.174	2.720	5.033	3.151	- 2.777	6.820	3.066
19	21.932	6.616	10.237	2.769	5.064	3.577	- 0.831	6.751	3.576
20	49.921	5.436	12.578	1.522	6.223	6.001	- 1.951	5.584	3.125
21	8.998	7.199	8.281	3.355	4.096	- 2.117	- 1.070	7.278	3.599
22	15.181	6.917	9.368	3.073	4.635	1.147	- 2.500	7.381	3.253
23	27.580	6.353	12.627	2.510	6.247	1.881	- 3.577	7.100	2.783
24	6.235	7.324	8.281	3.480	4.096	- 0.703	- 0.402	7.519	3.715

(Rappel des valeurs propres initiales : $\lambda_1(R) = 7.607$, $\lambda_2(R) = 3.763$)

L'importance de l'amélioration qu'apporte b_ℓ par rapport à m_ℓ ($\ell = 1,2$) n'est que la traduction de la propriété de croissance concave soulignée pour la fonction f introduite dans la démonstration de la proposition 2.6. Ainsi en ce qui concerne la première valeur propre et la borne b_1 on remarquera que les améliorations les plus importantes sont enregistrées pour les u.s. j pour lesquelles z_{j1}^2 est relativement faible par rapport à $||y_j||^2$ i.e. principalement les u.s. 5, 6, 7, 8, 13, 22, 23. L'amélioration apparait plus insignifiante pour des u.s. pour lesquelles z_{j1}^2 est proche de $||y_j||^2$ comme par exemple les u.s. 1, 2, 3, 4. On observe un phénomène analogue pour la deuxième valeur propre avec une amélioration plus importante (qu'apporte b_2 par rapport à m_2) pour les u.s. 1, 2, 3, 4, 19 et tout particulièrement pour l'u.s. 20 (z_{j2}^2 est petit par rapport à $||y_j||^2$) comparativement à l'amélioration moins nette pour des u.s. telles que les u.s. 5, 6, 7, 21 pour lesquelles la coordonnée z_{j2} joue un rôle important dans $||z_j|| = ||y_j||$.

Il est intéressant à titre indicatif d'avoir une idée de l'ordre de grandeur de la variation réelle des valeurs propres engendrée par la suppression des u.s. A cet effet le tableau qui suit donne les deux plus grandes valeurs propres $\lambda_\ell(\tilde{R})$, $\ell = 1,2$ correspondant à la suppression de chacune des u.s. 4, 5, 6, 9 et 20 en les mettant en rapport d'une part avec les valeurs propres initiales $\lambda_\ell(R)$, $\ell = 1,2$ et d'autre part avec les encadrements $[b_\ell, M_\ell]$, $\ell = 1,2$ qu'en donne le tableau précédent.

Valeurs propres initiales
$\lambda_1(R) = 7.607$
$\lambda_2(R) = 3.763$

u.s. supprimée	4	5	6	9	20
$\lambda_1(\tilde{R})$	7.223	7.924	7.863	7.682	6.907
b_1 ; M_1	6.356 ; 9.247	7.316 ; 8.990	7.568 ; 8.111	7.575 ; 7.910	5.584 ; 12.578
$\lambda_2(\tilde{R})$	4.019	3.336	3.449	3.738	4.139
b_2 ; M_2	3.620 ; 4.575	3.011 ; 4.447	3.288 ; 4.013	3.646 ; 3.913	3.125 ; 6.223

En ce qui concerne les encadrements eux-mêmes nous ne reprenons pas le détail des commentaires effectués en [3] et qui restent valables pour les nouvelles bornes. Notons simplement qu'un des intérêts pratiques des encadrements repose dans l'utilisation que l'on peut en faire en tant qu'aide à l'interprétation supplémentaire à même d'orienter l'utilisateur vers des u.s. jouant un rôle prépondérant sur les valeurs propres. Un encadrement d'amplitude faible autour de la valeur propre initiale indique immédiatement à l'utilisateur que l'u.s. supprimée ne joue qu'un rôle secondaire. A l'inverse des encadrements d'amplitude plus importante lui signale des u.s. pouvant éventuellement avoir une influence prépondérante. En ce sens il est souhaitable d'avoir des encadrements dont l'amplitude reflète de la manière la plus exacte possible la variation réelle des valeurs propres, d'où l'intérêt, nous semble-t-il, de l'amélioration qu'apporte les bornes b_ℓ par rapport à m_ℓ. On remarquera ainsi que l'amplitude relativement faible de l'encadrement $[b_1, M_1]$ pour la suppression de l'u.s. 5 et de $[b_2, M_2]$ pour l'u.s. 4 indique que ces u.s. ne jouent qu'un rôle secondaire sur la première et la deuxième valeur propre respectivement alors que l'amplitude des encadrements $[m_1, M_1]$ pour l'u.s. 5 et $[m_2, M_2]$ pour l'u.s. 4 ne permettaient pas d'arriver à cette conclusion. A l'inverse l'amplitude plus grande de $[b_1, M_1]$ pour la suppression de l'u.s. 20 et de $[b_2, M_2]$ pour celle des u.s. 5 et 20 (amplitude liée à une faible amélioration de b_ℓ par rapport à m_ℓ) signale aussi bien ces u.s. au statisticien comme étant susceptible d'influencer de manière sensible les inerties de certains axes. Les résultats du tableau précédent confirment alors le phénomène que laissait pressentir, en l'amplifiant, les encadrements, à savoir une variation très sensible de la première valeur propre lorsqu'on supprime l'u.s. 20 ($\lambda_1(R) - \lambda_1(\tilde{R}) = 0.699$) et une variation également non négligeable de la deuxième valeur propre aussi bien pour la suppression de l'u.s. 5 que de l'u.s. 20.

Remarque : Il est nécessaire que la condition :
$\lambda_{\ell-1}(R) - \lambda_\ell(R) > - \beta_j \left[||z_j||^2 - z_{j\ell}^2\right]$ soit vérifiée pour pouvoir appliquer la formulation (4.6) à la ℓ-ième valeur propre $\lambda_\ell(\tilde{R})$. Cette condition assure que le ℓ-ième valeur propre de $\lambda_\ell(R)$ est suffisamment séparée de $\lambda_{\ell-1}(R)$ qui la précède, par rapport à l'amplitude de la perturbation qui est caractérisée par $- \beta_j \left[||z_j||^2 - z_{j\ell}^2\right]$.

Nous avons vu que cette condition est vérifiée pour $\ell = 2$ c'est-à-dire la deuxième valeur propre. Il est facile de vérifier qu'elle est encore remplie pour $\ell = 3$ quelle que soit l'u.s. j que l'on se propose de supprimer puisqu'on a :

$$\lambda_2(R) - \lambda_3(R) > - \beta_j \left[||z_j||^2 - z_{j3}^2 \right] \iff 51.875 > \left[||y_j||^2 - z_{j3}^2 \right]$$

et $\max\limits_{\substack{j=1,..,24 \\ j \neq 17}} ||y_j||^2 = ||y_{20}||^2 = 49.921.$

Il est donc encore possible d'appliquer l'inégalité (4.6) à la valeur propre $\lambda_3(\tilde{R})$ dans tous les cas.

En ce qui concerne l'influence de la suppression d'une u.s. sur la quatrième valeur propre on a :

$$\lambda_3(R) - \lambda_4(R) > - \beta_j \left[||z_j||^2 - z_{j4}^2 \right] \iff 8.884 > \left[||y_j||^2 - z_{j4}^2 \right]$$

Cette condition apparait ici beaucoup plus restrictive et elle n'est plus remplie lorsqu'on se propose par exemple de supprimer l'u.s. 8 puisque l'on a avec $j = 8$

$||z_j||^2 = ||y_j||^2 = 15.907$, $z_{j1} = 1.500$, $z_{j2} = 2.925$, $z_{j3} = - 0.635$

$z_{j4} = - 0.151$ et donc :

$$||y_j||^2 - z_{j4}^2 = 15.884 > 8.884$$

Nous allons voir que s'il n'est donc plus légitime d'utiliser la relation (4.6), il reste toutefois encore possible d'améliorer la borne inférieure de l'encadrement (4.4) en se référant à la remarque 3 du paragraphe 3.2 qui transposée au cadre de la matrice de corrélation conduit pour une diminution de poids ($\beta_j < 0$) à l'inégalité

$$\lambda_\ell(\tilde{R}) \geqslant \left\{ \lambda_\ell(R) + \frac{1}{2} \beta_j \left[(z_{j\ell}^2 + z_{jk}^2) + \sqrt{(z_{j\ell}^2 + z_{jk}^2) + 4(z_{j\ell}^2 + z_{jk}^2)(||y_j||^2 - (z_{j\ell}^2 + z_{jk}^2))} \right] \right\} \left\{ \min_{k'} (1 + \beta_j y_{jk'}^2)^{-1} \right\}$$

si lorsque $\ell \neq 1$ on a

$$\lambda_{\ell-1}(R) - \lambda_\ell(R) > - \beta_j \left[||y_j||^2 - (z_{j\ell}^2 + z_{jk}^2) \right]$$

Dans le cadre de la suppression de l'u.s. 8 en prenant $\ell = 4$ et k = 2 on constate que la condition
$\lambda_3(R) - \lambda_4(R) > - \beta_j \left[||y_j||^2 - (z_{j4}^2 + z_{j2}^2)\right]$ est effectivement remplie puisqu'on a $\lambda_3(R) = 1.405$, $\lambda_4(R) = 1.001$. Il en découle

$$\lambda_3(R) - \lambda_4(R) = 0.404 \quad \text{et} \quad - \beta_j \left[||y_j||^2 - (z_{j4}^2 + z_{j2}^2)\right] = 0.333$$

On peut alors utiliser l'inégalité qui vient d'être présentée et qui donne :

$$\lambda_4(\tilde{R}) \geqslant 0.398$$

Ce résultat est à comparer avec ce que donne l'expression (4.4) à savoir :

$$\lambda_4(\tilde{R}) \geqslant 0.279$$

L'amélioration observée est due au fait que l'on a :

$$\frac{1}{2}\, \beta_j \left[(z_{j\ell}^2+z_{jk}^2) + \sqrt{(z_{j\ell}^2+z_{jk}^2)^2 + 4(z_{j\ell}^2+z_{jk}^2)(||y_j||^2 - (z_{j\ell}^2+z_{jk}^2))}\right] = - 0.605$$

alors que :

$$\beta_j \; ||y_j||^2 = - 0.787$$

En conclusion on voit donc que la relation (4.6) s'applique ici aux trois plus grandes valeurs propres (qui représentent 79.8 % de l'inertie). Cette constatation est encourageante si l'on souligne que nous nous sommes placés dans un exemple plutôt défavorable de par le faible nombre des u.s. Ceci induit un coefficient β_j relativement grand en valeur absolue (par rapport à des tableaux plus importants) et rend la condition :

$$\lambda_{\ell-1}(R) - \lambda_\ell(R) > - \beta_j \left[||y_j||^2 - z_{j\ell}^2\right]$$

plus restrictive. Lorsque cette condition n'est pas vérifiée (ce qui arrive pour les plus petites valeurs propres) les développements de la dernière remarque montrent qu'il est encore possible d'améliorer l'encadrement (4.4) avec des inégalités faisant intervenir deux coordonnées.

En fait la proposition 2.6 sous entend tout un ensemble d'inégalités avec un nombre de coordonnées pouvant varier entre 1 et p. On note que lorsqu'on augmente le nombre de coordonnées intervenant dans les inégalités ces dernières amènent une amélioration de moins en moins sensible (effective cependant tant qu'il n'y a pas plus de (p-1) coordonnées) mais que simultanément leurs conditions d'application deviennent aussi de moins en moins restrictives. Il est ainsi pratiquement toujours possible d'apporter une amélioration en utilisant l'une ou l'autre, même lorsque l'on s'intéresse à une valeur propre initialement proche des valeurs propres voisines.

BIBLIOGRAPHIE

[1] ANDERSON, T.W., "Some inequalities on characteristic roots of matrices", Biometrika, 1963, Vol. 50, pp. 522-524.

[2] BENASSENI, J., Une contribution à l'étude de la stabilité en analyse factorielle, Thèse de 3ème cycle, U.S.T.L. Montpellier, 1984.

[3] BENASSENI, J., "Influence des poids des unités statistiques sur les valeurs propres en analyse en composantes principales", Revue de Statistique Appliquée, 1985, Vol. 13, n° 4, pp. 41-55.

[4] BENASSENI, J., "Stabilité du pouvoir discriminant des facteurs par rapport à des perturbations des données en analyse linéaire discriminante", Statistique et Analyse des Données, 1985, Vol. 10, n° 2, pp. 1-28.

[5] ESCOFIER, B., Stabilité et approximation en analyse factorielle, Thèse de Doctorat d'Etat es Sciences, Université Pierre et Marie Curie, Paris, 1979.

[6] PAGES, J.P., CAILLIEZ, F., Introduction à l'Analyse des Données, SMASH, 1976.

[7] WILKINSON, J.H., The algebraic eigenvalue problem, Clarendon Press, Oxford, 1965.