DATA ANALYSIS AND INFORMATICS, III
E. Diday et al. (editors)



# APPROXIMATIONS D'APPLICATIONS LINEAIRES ET ANALYSE EN COMPOSANTES PRINCIPALES

R. SABATIER - Y. JAN - Y. ESCOUFIER

E.N.S.A.M. - I.N.R.A. - U.S.T.L.
UNITE DE BIOMETRIE
9, Place Pierre Viala
34060 MONTPELLIER CEDEX
FRANCE

Pour expliciter un ensemble de critères d'approximation optimisés par les solutions de l'Analyse en Composantes Principales, on rappelle des résultats sur les valeurs singulières d'une application linéaire d'un espace euclidien dans un autre, puis des théorèmes d'approximation concernant ces valeurs singulières. On montre alors que les normes unitairement invariantes sont des fonctions croissantes de ces valeurs, ce qui permet d'obtenir les résultats souhaités.

## INTRODUCTION

Réaliser une Analyse en Composantes Principales (ACP) consiste à substituer à un tableau de données un ensemble de graphiques et de nombres choisis pour mettre en évidence les ressemblances entre individus et les liaisons entre variables, ressemblances et liaisons qui sont contenues dans le tableau initial mais restent inaccessibles par une lecture directe.

Dans une présentation de l'Analyse des Données popularisée en France par F. CAILLIEZ et J.P. PAGES (1) au tableau T des données contenant les mesures de p variables quantitatives faites sur n individus sont associés :

- une forme bilinéaire définie positive Q sur $R^p$ permettant le calcul des distances entre les individus ;
- une forme diagonale positive D sur $R^n$ nécessaire au calcul des liaisons entre les variables ;
- le tableau X, n x p, des données centrées pour les poids définis par la diagonale de D.

Du triplet (X, Q, D) se déduisent alors l'opérateur $V = {}^tXDX$ de $R^p$ qui décrit les covariances, c'est-à-dire les indices de liaisons entre les variables et $W = XQ{}^tX$ de $R^n$ qui décrit les produits scalaires entre individus. L'ACP fournit p vecteurs, dans un espace de dimension faible dont les normes et les angles mutuels sont des approximations des éléments de V. Elle fournit de même des approximations des éléments de W.

Le but de cette contribution est de souligner que les approximations ainsi obtenues sont les meilleures au sens de critères très nombreux et pas seulement au sens des critères usuellement choisis pour introduire la méthode.

Pour atteindre cet objectif, on s'appuiera sur la notion de valeurs singulières d'une application linéaire. On rappellera des résultats d'approximation les concernant, ce qui permettra d'établir que les approximations obtenues sont optimales pour toutes les fonctions croissantes de ces valeurs. On terminera par une étude des normes unitairement invariantes qui conduit à expliciter une gamme de critères simples adaptés à l'appréciation de la qualité des approximations obtenues.

Pour conserver au texte une longueur raisonnable la plupart des démonstrations sont omises dans la première moitié du texte soit parce que les résultats concernés sont jugés suffisamment courants dans le contexte de l'Analyse des Données, soit à cause de leur simplicité.

VALEURS SINGULIERES D'UNE APPLICATION DE L(E,F)

On considère deux espaces vectoriels réels E et F de dimension respective p et q et les espaces euclidiens (E,M) et(F,N) obtenus en les munissant respectivement des formes bilinéaires symétriques définies positives M et N.

Soit L(E,F) l'ensemble des applications linéaires de E dans F et $A \in L(E,F)$.

Définition 1 : On appelle application adjointe de A, l'application $A^* \in L(F,E)$ définie par :

$\forall (x,y) \in E \times F : N(Ax,y) = M(x,A^*y)$

Définition 2 : Si $B \in L(E,E)$, B est appelé un opérateur de E. Il est dit M-symétrique si $B^* = B$.

Définition 3 : Un opérateur M-symétrique de E et idempotent est un M-projecteur.

Propriété 1 : Si ${}^tA$ est l'application transposée de A alors $A^* = M^{-1}\,{}^tA\,N$

Propriété 2 : Si $B \in L(E,E)$ est M-symétrique alors ${}^t(MB) = MB$.

Soit B un opérateur M-symétrique de E. Notons $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ ses valeurs propres non nécessairement distinctes et $\{\phi_i \; ; \; i=1, \dots p\}$ un ensemble de vecteurs propres associés de M-norme unité ; les vecteurs propres associés à des valeurs propres distinctes non nulles sont M-orthogonaux entre eux ; ils peuvent être choisis ainsi lorsqu'ils sont associés à des valeurs propres égales ou nulles. L'ensemble de ces vecteurs propres constitue un base M-orthogonale de E.

Propriété 3 : Soit $P_i \in L(E,E)$ défini par :

$\forall x \in E,\; P_i(x) = M(\phi_i,x)\phi_i$

Alors $B = \sum_{i=1}^{p} \lambda_i P_i$

Remarque : En identifiant M à l'application linéaire de E dans $E^*$ qu'elle permet de définir on peut écrire $M(\phi_i,x) = M(\phi_i)(x)$ ce qui permet d'adopter la notation $P_i = \phi_i \,.\, M(\phi_i)$

Propriété 4 : Les $P_i$ sont des M-projecteurs de E.

Conséquences : - $\sum_{i=1}^{p} P_i = I_E$ (Identité de E)

- Si r est le nombre de valeurs propres non nulles de B comptées avec leur multiplicité,

$\sum_{i=1}^{r} P_i = \mathcal{P}_r$ (projecteur sur Im(B))

$\sum_{i=r+1}^{p} P_i = \mathcal{P}_{p-r}$ (projecteur sur Ker(B))



Soit $A \in L(E,F)$ ; on considère $A^*A \in L(E,E)$

Propriété 5 : $A^*A$ est M-symétrique et ses valeurs propres non nulles sont positives.

Définition 4 : Notons $\lambda_1 \geq \lambda_r \dots \geq \lambda_r$ les r valeurs propres non nulles de $A^*A$. On a alors $A^*A = \sum_{i=1}^{r} \lambda_i P_i$

On appelle module de A et on note $|A|$, l'opérateur $|A| = \sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i$

Propriété 6 : Si B est un opérateur M-symétrique de L(E,E) à valeurs propres positives, $|B| = B$.

Proposition 7 : Pour tout $A \in L(E,F)$, il existe $U \in L(E,F)$ unique telle que $A = U|A|$

Démonstration : On remarque d'abord que $\mathrm{Ker}(A) = \mathrm{Ker}(A^*A)$. En effet

$Ay = 0 \Rightarrow A^*Ay = 0$ et $A^*Ay = 0 \Rightarrow N(y,A^*Ay) = 0 \Rightarrow N(Ay,Ay) = 0$

d'où $Ay = 0$

Soit alors $(P_i \; ; \; i=1, \dots, p)$ l'ensemble des projecteurs associés à une base orthonormée de vecteurs propres de $A^*A$ et $(P_i \; ; \; i=1, \dots, r)$ l'ensemble de ceux qui sont associés à des valeurs propres non nulles notées $(\lambda_i \; ; \; i=1, \dots, r)$.

On a : $A = A(\sum_{i=1}^{p} P_i) = A(\sum_{i=1}^{r} P_i) = \sum_{i=1}^{r} A P_i$

$= (\sum_{i=1}^{r} \frac{1}{\sqrt{\lambda_i}} A P_i)\,(\sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i)$

L'unicité découle de l'égalité $\mathrm{Ker}(U) = \mathrm{Ker}(A)$

Proposition 8 : $U U^* = \mathcal{P}_r$ (projecteur sur Im(A))

$I_E - U^*U = \mathcal{P}_{p-r}$ (projecteur sur Ker(A))

Définition 5 : On appelle valeurs singulières de $A \in L(E,F)$ les valeurs propres de $|A|$

Notations : Si $|A| = \sum_{i=1}^{r} \sqrt{\lambda_i}\, P_i$, on notera l'ensemble des valeurs singulières de A par $s(A) = (\sqrt{\lambda_1}, \dots \sqrt{\lambda_r}) = (s_1(A), \dots, s_r(A))$.

On remarquera que $s_k(A) = \left[s_k(A^*A)\right]^{1/2} = \left[s_k(AA^*)\right]^{1/2}$ et comme $(A^*)^* = A$ on se permettra par la suite de supposer $p \leqslant q$. S'il en était autrement, on travaillerait avec $A^*$. On pourra alors sans ambiguité écrire $s(A) = (\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_r}, 0, \ldots, 0) \in R_+^p$.

UN THEOREME D'APPROXIMATION CONCERNANT LES VALEURS SINGULIERES DE $A \in L(E,F)$

Les premiers résultats de ce paragraphe concernent les propriétés extrémales des éléments propres de $B = \sum_{i=1}^{p} \lambda_i P_i \in L(E,F)$. La propriété 13 énonce le résultat central pour notre point de vue.

Propriété 9 :

$$\forall x \in E, \quad x \neq 0 \quad : \quad \frac{M(Bx,x)}{M(x,x)} \leqslant \frac{M(B\phi_1,\phi_1)}{M(\phi_1,\phi_1)} = \lambda_1$$

$$\forall x \in E, \quad x \neq 0 \quad : \quad \frac{M(Bx,x)}{M(x,x)} \geqslant \frac{M(B\phi_p,\phi_p)}{M(\phi_p,\phi_p)} = \lambda_p$$

Si pour $1 \leqslant k \leqslant p$, $E_k(B)$ est le sous-espace vectoriel de E engendré par $\phi_1, \ldots, \phi_k$, et $E_k^{\perp}(B)$ son orthogonal, on peut énoncer.

Propriété 10 :

$$\forall x \in E_k^{\perp}(B),\ x \neq 0 \quad \frac{M(Bx,x)}{M(x,x)} \leqslant \frac{M(B\phi_{k+1},\phi_{k+1})}{M(\phi_{k+1},\phi_{k+1})} = \lambda_{k+1}$$

$$\forall x \in E_{k+1}(B)\ x \neq 0 \quad \frac{M(Bx,x)}{M(x,x)} \geqslant \frac{M(B\phi_{k+1},\phi_{k+1})}{M(\phi_{k+1},\phi_{k+1})} = \lambda_{k+1}$$

Si $E \geqslant a$ (respectivement $E \leqslant a$ ; $E = a$) désigne l'ensemble des sous-espaces vectoriels de E de dimensions supérieures ou égales à a (respectivement inférieure ou égale à a ; égale à a) on a :

Propriété 11 : (Théorème de Courant-Fisher)

$\forall k = 1, \ldots, p-1$

$$\inf_{H \in E \geqslant p-k} \ \sup_{\substack{x \in H \\ x \neq 0}} \frac{M(Bx,x)}{M(x,x)} = \sup_{\substack{x \in E_k^{\perp}(B) \\ x \neq 0}} \frac{M(Bx,x)}{M(x,x)} = \lambda_{k+1}$$

$$\sup_{H \in E \geqslant p-k} \ \inf_{\substack{x \in H \\ x \neq 0}} \frac{M(Bx,x)}{M(x,x)} = \inf_{\substack{x \in E_{p-k}(B) \\ x \neq 0}} \frac{M(Bx,x)}{M(x,x)} = \lambda_{p-k}$$



Propriété 12 : (Théorème de Sturm)

Si P est un M-projecteur de E de rang p-k tel que $s(P B P) = (\mu_1, \ldots, \mu_p)$, on a : $\forall i \in (1, \ldots, p-k) \quad \lambda_{i+k} \leqslant \mu_i \leqslant \lambda_i$

$\forall i \in (p-k+1, \ldots, p) \quad \mu_i = 0$

Remarque : On trouve un énoncé de la réciproque de ce théorème dans (5) et une démonstration dans (3)

Propriété 13 : Soit $A \in L(E,F)$ de rang r telle que $A^*A = \sum_{i=1}^{r} \lambda_i P_i$ et $A = U|A|$.

Soit $A_k \in L(E,F)$ de rang $k \leqslant r$

Alors 1) $s_i(A-A_k) \geqslant s_{i+k}(A)$ si $i+k \leqslant r$

$s_i(A-A_k) \geqslant 0$ si $i+k > r$

2) Les égalités dans 1 sont atteintes pour $\tilde{A}_k = U(\sum_{i=1}^{k} \sqrt{\lambda_i}\, P_i)$

Démonstration :

1) Soit $A_k = V|A_k|$ la décomposition de $A_k$ que permet la proposition 7.

Posons $B = (I_E - V^*V)(A-A_k)^*(A-A_k)(I_E-V^*V)$

On a :

$$s_i^2(A-A_k) = \lambda_i((A-A_k)^*(A-A_k)) \geqslant \lambda_i(B) = \sup_{\substack{x \in E_{i-1}^{\perp}(B) \\ x' \neq 0}} \left\{\frac{M(Bx,x)}{M(x,x)}\right\}$$

Posons alors $\tilde{B} = (I_E-V^*V)A^*A(I_E-V^*V)$

On remarque que B et $\tilde{B}$ coïncident sur $\mathrm{Ker}(A_k)$ si bien que :

$$s_i^2(A-k) \geqslant \sup_{\substack{x \in E_{i-1}^{\perp}(B) \\ x \neq 0}} \left\{\frac{M(\tilde{B}x,x)}{M(x,x)}\right\}$$

$$\geqslant \inf_{H \in E \geqslant p-(i-1)} \ \sup_{\substack{x \in H \\ x \neq 0}} \left\{\frac{M(\tilde{B}x,x)}{M(x,x)}\right\}$$

$$= \lambda_i((I_E-V^*V)A^*A(I_E-V^*V))$$

D'après la proposition 12, on a donc :

pour $i \leqslant r-k \quad \lambda_{i+k}(A^*A) \leqslant s_{i+k}^2(A-A_k)$

et pour $i > r-k$ $\qquad s_i^2(A-A_k) \geq 0$

2) Le résultat s'obtient en remarquant que $A-\tilde{A}_k = A\left(\sum_{i=k+1}^{p} P_i\right)$

On note $H_k^+ = \{\underline{a} = (a_1, \ldots, a_k) \in R_+^k / a_1 \geq a_2 \ldots \geq a_k\}$

et sur $H_k^+$, on définit la relation d'ordre partiel $\forall \underline{a} \in H_k^+$, $\forall \underline{b} \in H_k^+$

$$\{\underline{a} \leq \underline{b}\} \Longleftrightarrow \{a_i \leq b_i, \text{ pour tout } i=1, \ldots, k\}$$

<u>Définition 6</u> : On dit que la fonction $f : R^k \to R$ est croissante sur $H_k^+$ pour $\leq$ si et seulement si $\{\forall \underline{a} \in H_k^+, \forall \underline{b} \in H_k^+, \underline{a} \leq \underline{b}\} \Rightarrow \{f(\underline{a}) \leq f(\underline{b})\}$

<u>Propriété 14</u> : Soit f une fonction croissante pour $\leq$ sur $H_p^+$. Pour toute application $A \in L(E,F)$, on pose

$$\phi(A) = f(s(A))$$

Alors $$\min_{\substack{A_k \in L(E,F) \\ rg(A_k) = k}} \phi(A-A_k) = \phi(A-\tilde{A}_k) = \phi\left(A\left(\sum_{i=k+1}^{p} P_i\right)\right)$$

La démonstration découle directement de la propriété 13 qui a établit que pour toute application

$$A_k \in L(E,F),\ s(A-A_k) \geq s(A-\tilde{A}_k)$$

Remarque : On peut ici faire un inventaire des fonctions f croissantes pour $\leq$ sur $H_p^+$. Y. JAN, dans un travail récent ( 6 ) démontre l'équivalence des propositions suivantes :

1) $\forall (a_1, \ldots, a_p) \in H_p^+$, $\forall (b_1, \ldots, b_p) \in H_p^+$

$$(a_1, \ldots, a_p) \leq (b_1, \ldots, b_p) \Rightarrow f(a_1, \ldots, a_{p-1}, 0) \leq f(b_1, \ldots, b_p)$$

2) $\forall k \in 1, \ldots, p$

$$\min_{\substack{A_k \in L(E,F) \\ rg(A_k) = k}} \phi(A-A_k) = \phi(A-\tilde{A}_k)$$

Le paragraphe suivant s'intéresse à la classe particulière des fonctions associées aux normes unitairement invariantes sur L(E,F)



<u>Définition 7</u> : Une norme notée $\|\ \|$ sur L(E,F) est dite unitairement invariante si et seulement si pour tout $A \in L(E,F)$

tout $U \in L(E,E)$ tel que $U^*U = I_E$

tout $V \in L(F,F)$ tel que $V^*V = I_F$

on a $\|A\| = \|UAV^*\|$

<u>Propriété 15</u> : L'application $\|\ \|_t$ de L(E,F) dans $R^+$ définie par

$$\|A\|_t = \left[\sum_{i=1}^{p} s_i^2(A)\right]^{1/2}$$ est une norme unitairement invariante sur L(E,F).

<u>Remarque</u> : Cette norme est en fait un cas très particulier car elle découle du produit scalaire :

$$\left.\begin{array}{l} A_1 \in L(E,F) \\ A_2 \in L(E,F) \end{array}\right\} \longrightarrow \left[\sum_{i=1}^{p} s_i(A^*B)\right]^{1/2}$$

Elle nous sera utile dans le paragraphe suivant.

<u>Propriété 16</u> : Toute norme définie sur L(E,F) à partir des valeurs singulières est une norme unitairement invariante.

<u>Démonstration</u> :

D'après les paragraphes précédents

$$s_i^2(UA) = s_i((UA)^*(UA)) = s_i(A^*U^*UA) = s_i(A^*A) = s_i^2(A)$$

et de même

$$s_i^2(AV^*) = s_i(AV^*(AV^*)^*) = s_i(AV^*VA^*) = s_i(AA^*) = s_i^2(A)$$

Il en découle que $s_i^2(UAV^*) = s_i^2(A)$ ce qui établit le résultat.

<u>Propriété 17</u> : Si A et $B \in L(E,F)$ ont les mêmes valeurs singulières alors pour toute norme unitairement invariante $\|A\| = \|B\|$

Soit $\{\phi_1, \ldots, \phi_p\}$ une base de E formée des vecteurs propres M-orthonormés de $A^*A$ éventuellement complétée par des vecteurs de Ker(A). D'après la propriété 3 on peut écrire :

$$\forall x \in E \quad A(x) = \sum_{i=1}^{p} M(\phi_i, x)\, A(\phi_i)$$

D'où l'on déduit que pour tout opérateur

$V \in L(E,E)$ et $U \in L(F,F)$

$$\forall x \in E \quad UAV^*(x) = \sum_{i=1}^{p} M(V(\phi_i), x)U(A(\phi_i))$$

On peut remarquer également que les $A(\phi_i)$ forment une famille orthogonale de vecteurs propres de $AA^*$. On peut alors construire une base de F, N-orthonormale $(\psi_1, \ldots, \psi_q)$ dont les r premiers vecteurs seront les $A(\phi_i)/s_i(A)$

$B \in L(E,F)$ ayant les mêmes valeurs singulières que A, peuvent lui être associées de la même manière une base $(\phi'_1, \ldots, \phi'_p)$ de E et une base

$(\psi'_1, \ldots, \psi'_q)$ de F

Prenons pour U l'application unitaire de F définie par $(U(\psi_i) = \psi'_i$ ; $i=1, \ldots, q)$

et pour V, l'application unitaire de E définie par $(V(\phi_i) = \phi'_i)$ ; $i=1, \ldots, p)$.

On a alors pour tout $x \in E$

$$UAV^*(x) = \sum_{i=1}^{p} M(V(\phi_i),x)U(A(\phi_i))$$

$$= \sum_{i=1}^{r} M(V(\phi_i),x)U(A(\phi_i))$$

$$= \sum_{i=1}^{r} s_i(A)M(V(\phi_i),x)U(\psi_i)$$

$$= \sum_{i=1}^{r} s_i(B)M(\phi'_i,x)\psi'_i$$

$$= \sum_{i=1}^{r} M(\phi'_i,x)B(\phi'_i)$$

$$= \sum_{i=1}^{p} M(\phi'_i,x)B(\phi'_i)$$

$$= B(x)$$

Il en découle pour toute norme unitairement invariante $||B|| = ||UAV^*|| = ||A||$

Propriété 18 : Il existe une application $K \in L(E,F)$ telle que pour tout $B \in L(E,E)$, $s(B) = s(KB)$

Démonstration :

$$s_i^2(B) = s_i(B^*B) = s_i(M^{-1t}BMB)$$

et $$s_i^2(KB) = s_i((KB)^*KB) = s_i(M^{-1t}B^tKNKB)$$

Le problème est donc de savoir s'il existe K telle que ${}^tKNK = M$ ce qui est réalisé puisque N et M sont symétriques définies positives avec rang (N) = p $\leqslant$ rang (M) = q.



Propriété 19 : A toute norme $|| \, ||_1$ unitairement invariante sur L(E,F) correspond une norme $|| \, ||_2$ unitairement invariante sur L(E,E) et $||A||_1 = || \, |A| \, ||_2$

Démonstration :

Si $|| \, ||_1$ est unitairement invariante sur L(E,F), il existe f telle que

$||A||_1 = f(s(A))$

Définissons alors sur L(E,E)

$||B||_2 = f(s(B))$

Il faut montrer que $|| \, ||_2$ est une norme. D'après la propriété 15 elle sera nécessairement unitairement invariante.
La démonstration découle simplement de $s(B) = s(KB)$ qui permet d'écrire

$||B||_2 = ||KB||_1$

et d'en déduire les propriétés de $|| \, ||_2$ à partir de celles de $|| \, ||_1$

L'égalité $s(A) = s(|A|)$ implique que $|| \, |A| \, ||_2 = ||KB||_2$ ce qui complète le résultat.

Propriété 20 : Si $|| \, ||_1$ est une norme unitairement invariante sur L(E,F), l'application f telle que $||A||_1 = f(s(A))$ est croissante pour $\leqslant$ sur $H_p^+$

Démonstration : On va montrer que le résultat est vrai pour les opérateurs positifs de L(E,E) (Théorème de Schatten). Le résultat précédent $||A||_1 = || \, |A| \, ||_2$ établira la propriété.

On considère B et C des opérateurs positifs de L(E,E).

$$B = |B| = \sum_{i=1}^{p} \lambda_i P_i$$

$$C = |C| \quad \sum_{i=1}^{p} \mu_i Q_i$$

On suppose $s(B) \leqslant s(C)$ et on considère

$$C_1 = \lambda_1 Q_1 + \sum_{i=2}^{p} \mu_i Q_i$$

Posons $\alpha = \dfrac{\lambda_1 + \mu_1}{2\mu_1} \leqslant 1$ On a $\lambda_1 = \alpha\mu_1 - (1-\alpha)\mu_1$

$$\text{D'où } C_1 = \alpha \left(\sum_{i=1}^{p} \mu_i Q_i\right) + (1-\alpha) \left(-\mu_1 Q_1 + \sum_{i=2}^{p} \mu_i Q_i\right)$$

$$||C_1||_2 \leqslant \alpha \left|\left|\sum_{i=1}^{p} \mu_i Q_i\right|\right|_2 + (1-\alpha) \left|\left| -\mu_1 Q_1 + \sum_{i=2}^{p} \mu_i Q_i\right|\right|_2$$

parce que $|| \, ||_2$ ne dépend que des valeurs singulières on peut écrire :

$$||C_1||_2 \leq \alpha\, ||C||_2 + (1-\alpha)||C||_2 = ||C||_2$$

On montrerait de même que pour $C_2 = \lambda_1 Q_1 + \lambda_2 Q_2 + \sum_{i=3}^{p} \mu_i Q_i$

On a $||C_2||_2 \leq ||C_1||_2 \leq ||C||_2$

et de proche en proche, on arrivera à $||B||_2 \leq ||C||_2$

Conséquence : Pour toute norme unitairement invariante $||\ ||_1$ sur L(E,F), on a, avec les notations de la proposition 14

$$\min_{\substack{A_k \in L(E,F) \\ rg(A_k) = k}} ||A-A_k||_1 = ||A-\tilde{A}_k||_1$$

EXPLICITATION DES RESULTATS DANS LE CONTEXTE DE L'ANALYSE DES DONNEES

Dans la présentation de l'Analyse des Données évoquée dans l'Introduction, au triplet (X,Q,D) est associé le schéma de dualité :

$$\begin{tikzcd} E \arrow[d, "Q"'] & & F^* \arrow[ll, "{}^tX"'] \\ E^* \arrow[rr, "X"'] & & F \arrow[u, "D"'] \end{tikzcd}$$

où Q et D formes bilinéaires symétriques définies sont identifiées aux applications linéaires de E dans $E^*$ et F dans $F^*$ qui leur sont associées.

Du schéma se déduisent les opérateurs

$$VQ = {}^tXDXQ \quad \text{de} \quad E$$
$$\text{et} \quad WD = XQ\,{}^tXD \quad \text{de} \quad F$$

L'Analyse en composantes principales fournit les axes principaux de E($\phi_i$ ; i=1, ..., p) vecteurs propres de VQ orthonormés pour Q et les axes principaux de F ($\psi_i$ ; i=1, ..., p) vecteurs propres de WD othornormés par D.

Le classique calcul des "formules de transition" montre que :

$$\forall i = 1, \ldots, p \qquad XQ\phi_i = \sqrt{\lambda_i}\,\psi_i$$
$$\text{et} \quad {}^tXD\psi_i = \sqrt{\lambda_i}\,\phi_i$$

Considérons les applications XQ $\in$ L(E,F)

$$(VQ)^{\alpha} = \sum_{i=1}^{p} \lambda_i^{\alpha} P_i = \sum_{i=1}^{p} \lambda_i^{\alpha}\, \phi_i\, Q(\phi_i) \in L(E,E) \ ; \ \alpha > 0$$



On a $(XQ)^* = {}^tXD$ et $(XQ)^*(XQ) = VQ$

$((VQ)^{\alpha})^* = (VQ)^{\alpha}$ et $((VQ)^{\alpha})^*(VQ)^{\alpha} = VQ^{2\alpha}$

Il découle alors des paragraphes précédents que :

$$\widetilde{(XQ)}_k = \sum_{i=1}^{k} XQP_i = \sum_{i=1}^{k} \sqrt{\lambda_i}\, \psi_i . Q(\phi_i)$$

$$\widetilde{((VQ)^{\alpha})}_k = \sum_{i=1}^{k} \lambda_i^{\alpha}\, \phi_i . Q(\phi_i)$$

et en particulier

$$||XQ - \widetilde{(XQ)}_k||_t^2 = \sum_{i=k+1}^{p} \lambda_i = \left(\sum_{i=1}^{p} \lambda_i\right)\left(1 - \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{p} \lambda_i}\right)$$

$$= ||XQ||_t^2 \left(1 - \frac{||\widetilde{(XQ)}_k||_t^2}{||XQ||_t^2}\right)$$

$$||(VQ)^{\alpha} - \widetilde{(VQ)}_k^{\alpha}||_t^2 = \sum_{i=k+1}^{p} \lambda_i^{2\alpha} = \left(\sum_{i=1}^{p} \lambda_i^{2\alpha}\right)\left(1 - \frac{\sum_{i=1}^{k} \lambda_i^{2\alpha}}{\sum_{i=1}^{p} \lambda_i^{2\alpha}}\right)$$

$$= ||(VQ)^{\alpha}||_t^2 \left(1 - \frac{||\widetilde{(VQ)}_k^{\alpha}||_t^2}{||(VQ)^{\alpha}||_t^2}\right)$$

Remarque : En notant que $((WD)^{\alpha/2}XQ)^* = (XQ)^*((WD)^{\alpha/2})^* = {}^tXD(WD)^{\alpha/2}$

on obtient que $((WD)^{\alpha/2}XQ)^*((WD)^{\alpha/2}XQ) = (VQ)^{\alpha+1}$ ce qui permet d'obtenir des formules analogues pour l'application $(WD)^{\alpha/2}XQ \in L(E,F)$ pour tout $\alpha > 0$.
On aurait des résultats semblables pour ${}^tXD$, WD et $(VQ)^{\alpha\,2}\,{}^tXD$.

Ces résultats donnent un sens précis aux critères habituels $\left(\sum_{i=1}^{k} \lambda_i \Big/ \sum_{i=1}^{p} \lambda_i\right)$ et $\left(\sum_{i=1}^{k} \lambda_i^2 \Big/ \sum_{i=1}^{p} \lambda_i^2\right)$. Ils mettent en évidence qu'ils ne sont que deux critères particuliers parmi une infinité de critères possibles ($\alpha > 0$). Leurs choix découlent bien sûr de leur signification pratique : reconstruction de XQ et ${}^tXD$, VQ et WD. D'autres critères d'approximation pourraient être exhibés en remplaçant $||\ ||_t$ par d'autres normes unitairement invariantes ou même d'autres fonctions croissantes des valeurs singulières des applications considérées.

On remarquera de plus que l'application de la propriété 12 à VQ et WD ouvrent d'autres possibilités pour définir des critères qui seront toujours optimisés par les solutions de l'ACP.

## BIBLIOGRAPHIE

[1] CAILLIEZ, F. et PAGES, J.P., Introduction à l'Analyse des Données. (SMASH, 9, rue Duban, 75 016 Paris, 1976).

[2] CHAO, W. CHEN, An optimal property of Principal Components. Communications in Statistics. 3 (10) (1974) 979-983.

[3] CROQUETTE, A., Quelques résultats synthétiques en Analyse des Données Multidimensionnelles. Optimalité et métriques à effets relationnels. Thèse 3e cycle. Univ. Paul Sabatier, Toulouse (1980).

[4] DARROCK,J.N., An optimal property of Principal Components, Ann. Math. Statist. 36 (1965) 1579-1582.

[5] GLAZMAN, I. and LIUBITCH, Y., Analyse linéaire dans les espaces de dimensions finies. (Edition de Moscou, 1972).

[6] JAN, Y., A propos des théorèmes d'approximation, Note technique, Unité de Biométrie, Montpellier (1982).

[7] KOBILINSKI, A., Décomposition de formes quadratiques en Analyse des Données. Thèse 3e cycle. Univ. Paris Sud (1979).

[8] KOBILINSKI, A., Ordre entre formes quadratiques. Application à l'optimalité des sous-espaces en Analyse des Données, R.S.A., Vol XXVII n° 1 (1979) 45-54.

[9] LEBART, L., MORINEAU, A. et FENELON, J.P., Traitement des Données Statistiques : Méthodes et Programmes (Dunod, Paris, 1979).

[10] OKAMOTO, M. and KANAZAWA, M., Minimization of eigen values of a matrix and optimality of principal components, Annals of Mathematic Statistics, Vol. 37, (1968) 1736-1746.

[11] OKAMOTO, M., Optimality of Principal Components in Multivariate Analysis II, (P.K. Krishnaiah Ed, Academic Press, 1969, 673-685).

[12] RAO, C.R., Matrix approximation and reduction of dimensionality on multivariate statistical analysis,in Multivariate Analysis V (North Holland Publishing Company (1980) 3-22).

[13] SABATIER, R., Approximations d'un tableau de données. Application à la reconstitution des paléoclimats. Thèse 3e cycle, Université des Sciences et Techniques du Languedoc, Montpellier (1983).

[14] SCHATTEN, N., Norms ideals of completely continuous operator (Springer Verlag, 1960).