

M. BENAYOUN

J.-P. BENZÉCRI

M. FROISSART

Sur la stabilité des sous-espaces principaux d'inertie sous un changement de métrique

Les cahiers de l'analyse des données, tome 13, n° 2 (1988),
p. 197-203

http://www.numdam.org/item?id=CAD_1988__13_2_197_0

© Les cahiers de l'analyse des données, Dunod, 1988, tous droits réservés.
L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

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

SUR LA STABILITÉ DES SOUS-ESPACES PRINCIPAUX D'INERTIE SOUS UN CHANGEMENT DE MÉTRIQUE

[STABILITÉ]

M. BENAYOUN*

J.-P. BENZÉCRI**

M. FROISSART*

NB Une première version du présent travail, due à M. B. et M. F., a été diffusée, sous la cote LPC 85-06, par le Laboratoire de Physique Corpusculaire du Collège de France. J.-P. B. a proposé de généraliser les résultats obtenus en les plaçant dans le cadre géométrique propre à l'analyse des données multidimensionnelles. La présente rédaction est due, pour l'essentiel, à M. F.

0 Recherche des directions principales et recherche des contraintes

L'analyse statistique d'une distribution multidimensionnelle conduit ordinairement à rechercher des directions principales, ou directions dans lesquelles l'échantillon analysé a ses plus grandes dimensions. Parfois, on s'intéresse à découvrir des contraintes liant les variables; en particulier, les contraintes linéaires, exactes ou approchées, se manifestent par l'existence de directions dans lesquelles le nuage de points a une épaisseur nulle ou négligeable.

Une telle analyse comporte nécessairement des comparaisons entre distances mesurées dans des directions différentes: ces comparaisons se fondent sur le choix d'un tenseur métrique. Diverses règles ont été proposées pour guider ce choix; mais, d'une part, aucune de ces règles ne nous paraît s'imposer absolument; et, d'autre part, leur application précise n'est même pas toujours possible.

(*) Laboratoire de Physique Corpusculaire du Collège de France.

(**) Université Pierre et Marie Curie.

On se propose ici d'examiner dans quelle mesure une incertitude sur la métrique peut rendre incertaine ou même illusoire la recherche des directions principales et des contraintes.

Il apparaîtra que, pouvu qu'il existe, pour le nuage considéré, un grand rapport entre les dimensions principales les plus grandes et celles qui les suivent, des variations sur la métrique d'un ordre de grandeur inférieur à celui de ce rapport seront sans conséquences pour la détermination des contraintes. Nous devons toutefois souligner que, si cette condition n'est pas vérifiée, la recherche des directions principales ne produit que des résultats dépourvus de signification.

1 Angle entre sous-espaces d'un espace euclidien E

Dans la section présente, et la suivante, nous poserons quelques notations et lemmes se référant à une seule structure métrique. On pourra donc supposer que E est l'espace numérique R^n muni de la métrique usuelle définie par la somme des carrés des différences des coordonnées. Un sous-espace vectoriel étant noté l, on désignera par $n(l)$ le sous-espace supplémentaire orthogonal à E, ie. le sous-espace formé des vecteurs orthogonaux à tout vecteur de l. On désignera par l° l'ensemble des vecteurs non-nuls de l; et de même par $n^\circ(l)$ l'ensemble des vecteurs non nuls de $n(l)$. On notera $\theta(x,y)$ l'angle, compris entre 0 et π , formé par deux vecteurs x et y de E.

1.1 Angle d'un sous-espace et d'un vecteur

Définition

L'angle entre un vecteur x et un sous-espace l est défini par

$$\theta(x,l) = \min\{\theta(x,y) \mid y \in l^\circ\} \leq \pi/2 \quad (1.1)$$

Si x n'est pas orthogonal à l, le minimum est atteint pour y colinéaire à la projection orthogonale de x sur l, et $\theta(x,l) < \pi/2$.

Si l_1, \dots, l_k sont des sous-espaces mutuellement orthogonaux, qui sous-tendent un sous-espace l, on a, par projection orthogonale de x sur les l_j :

$$\Sigma\{\cos^2\theta(x,l_j) \mid j = 1, \dots, k\} = \cos^2\theta(x,l) \quad (1.2)$$

En particulier, si $k=2$ et $l = E$, il vient:

$$\cos^2\theta(x,l) = \sin^2\theta(x,n(l)) \quad (1.3)$$

1.2 Angle de deux sous-espaces

Définition

L'angle entre deux sous-espaces l et l' avec $\dim l \leq \dim l'$ est :

$$\begin{aligned} \theta(l,l') &= \text{Max}\{\theta(x,l') \mid x \in l^\circ\} \\ &= \text{Max}\{\min\{\theta(x,y) \mid y \in l'^\circ\} \mid x \in l^\circ\} \end{aligned} \quad (1.4)$$

Propriété

$$\theta(l,l') = \theta(n(l'),n(l)) \quad (1.5)$$

En effet, par (1.4) et (1.3) on a:

$$\begin{aligned} \theta(l,l') &= \text{Max}\{(\pi/2) - \min\{\theta(x,y) \mid y \in n^\circ(l')\} \mid x \in l^\circ\} \\ &= \text{Max}\{(\pi/2) - \theta(x,y) \mid y \in n^\circ(l'), x \in l^\circ\} \end{aligned}$$

et par le chemin inverse:

$$\begin{aligned} \theta(l,l') &= \text{Max}\{(\pi/2) - \min\{\theta(x,y) \mid x \in l^\circ\} \mid y \in n^\circ(l')\} \\ &= \text{Max}\{\min\{\theta(x,y) \mid x \in n^\circ(l)\} \mid y \in n^\circ(l')\}. \end{aligned}$$

2 Inertie

NB Dans tout ce §, par hyperplan ou par sous-espace de dimension p , nous entendrons: hyperplan ou sous-espaces de dimension p passant par le centre de gravité du nuage des points x_i affectés des masses m_i , considéré dans un espace E de dimension n :

$$N(I) = \{ (x_i, m_i) \mid i \in I \}.$$

On pourra supposer que ce centre de gravité est l'origine, mais cela importe peu puisque nous ne faisons aucun calcul sur les coordonnées.

2.1 Définition de l'inertie

Il est d'usage de considérer l'inertie dans une *direction*, et en particulier dans celle des axes principaux d'inertie. Pour comparer diverses métriques, il est plus pratique de considérer l'inertie transverse à un hyperplan (sous-espace de dimension $n-1$). Ces notions se confondent en prenant l'hyperplan et la direction normale. Mais l'inertie transverse par rapport à un hyperplan h s'exprime directement en fonction de la *distance*:

$$\text{Itr}(h) = \Sigma \{ m_i d^2(x_i, h) \mid i \in I \} \quad (2.1)$$

où on a noté, comme de règle:

$$d(x, h) = \min \{ d(x, y) \mid y \in h \} \quad (2.2)$$

2.2 Éléments principaux

Les moments principaux afférents au nuage $N(I)$ sont les valeurs stationnaires de Itr . Nous les supposons ordonnées par valeurs non-croissantes:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \quad (2.3)$$

Nous désignerons par e_1, e_2, \dots, e_n une suite de vecteurs propres (orthonormés), tels que $\text{Itr}(n(e_j)) = \lambda_j$, et par l_k le sous-espace engendré par e_1, \dots, e_k .

2.3 Expression de l'inertie transversale en fonction des angles

Soit $x \in n^\circ(h)$; projetons x sur les e_j , il vient:

$$\begin{aligned} \text{Itr}(h) &= \Sigma \{ \lambda_j \cos^2 \theta(x, e_j) \mid j = 1, \dots, n \} \\ &= \Sigma \{ \lambda_j \sin^2 \theta(e_j, h) \mid j = 1, \dots, n \} \end{aligned} \quad (2.4)$$

En vertu de (1.2) et de (1.3), on peut écrire:

$$\Sigma \{ \cos^2 \theta(e_j, x) \mid j = 1, \dots, p \} = \sin^2 \theta(l_p, h) \quad (2.5)$$

et donc

$$\Sigma\{\sin^2\theta(e_j, x) \mid j = p+1, \dots, n\} = \cos^2\theta(l_p, h) \quad (2.6)$$

2.4 Inégalités fondamentales

En faisant la moyenne sur les p premiers termes et les $n-p$ derniers dans (2.4), et utilisant (2.5), (2.6), il vient:

$$\text{Itr}(h) = \alpha \sin^2\theta(l_p, h) + \beta \cos^2\theta(l_p, h) \quad (2.7)$$

avec $\lambda_1 \geq \alpha \geq \lambda_p$ et $\lambda_{p+1} \geq \beta \geq \lambda_n$.

On en tire, d'une part :

$$\text{Itr}(h) \geq \lambda_p \sin^2\theta(l_p, h) \quad (2.8)$$

qui nous servira à borner les angles, et d'autre part une formule variationnelle pour les λ directement à partir de Itr :

$$\begin{aligned} \lambda_{p+1} &= \min \{ \text{Max} \{ \text{Itr}(h) \mid h \supset l \} \mid l \in L_p \} \\ &= \text{Max} \{ \text{Itr}(h) \mid h \supset l_p \} \end{aligned} \quad (2.9)$$

où L_p est l'ensemble des sous-espaces de dimension p de l'espace ambiant E . En effet, il suffit de prendre $h = l \oplus n(l_{p+1})$; $n(h)$ est orthogonal à e_j , pour $j=p+1, \dots, n$, et $\beta=l_{p+1}$ dans (2.8). Donc:

$$\text{Max} \{ \text{Itr}(h) \mid h \supset l \} \geq \alpha \sin^2\theta(l_p, h) + \lambda_{p+1} \cos^2\theta(l_p, h) \quad (2.10)$$

Le membre de droite est minimisé pour $\theta(l_p, h) = 0$, c'est-à-dire $l = l_p$. On a effectivement, dans ce cas, $\lambda_{p+1} = \text{Itr}(l_p)$; c.q.f.d.

3 Comparaisons entre métriques

Nous supposons que nous avons sur E deux métriques euclidiennes A et B , satisfaisant aux inégalités:

$$\forall x, y \in E: k d_A^2(x, y) \leq d_B^2(x, y) \leq K d_A^2(x, y) \quad (3.1)$$

où k et K sont deux constantes convenables. Les éléments principaux d'inertie du nuage $N(I)$ ainsi que les sous-espaces orthogonaux $n(I)$, afférents aux deux métriques seront, dans la suite distingués par les deux indices A et B .

3.1 Résultat sur les moments principaux

Théorème Les conditions (3.1) entraînent des relations analogues pour les moments principaux d'inertie:

$$k \lambda_{A_p} \leq \lambda_{B_p} \leq K \lambda_{A_p} \quad (3.2)$$

Il découle en effet de la définition (2.1) de I_{tr} et de (3.1) que:

$$k I_{tr_A}(h) \leq I_{tr_B}(h) \leq K I_{tr_A}(h) \quad (3.3)$$

La relation de minimax (2.9) conserve des inégalités partout vérifiées, c.q.f.d.

3.2 Résultats sur les sous-espaces de composantes principales et de contraintes

Théorème Les conditions (3.1) entraînent les majorations suivantes:

$$\sin^2 \theta_A(l_{A_p}, l_{B_p}) \leq (K/k) (\lambda_{A_{p+1}} / \lambda_{A_p}) \quad (3.4)$$

De (2.8) et (3.3) on tire qu'est vérifié pour tout hyperplan h :

$$\lambda_{A_p} \sin^2 \theta_A(l_{A_p}, h) \leq I_{tr_A}(h) \leq (1/k) I_{tr_B}(h)$$

En particulier, par (2.9):

$$\text{Max} \{ \lambda_{A_p} \sin^2 \theta_A(l_{A_p}, h) \mid h \supset l_{B_p} \} \leq (1/k) \lambda_{B_{p+1}}$$

Par (1.5) ceci devient:

$$\lambda_{A_p} \text{Max} \{ \sin^2 \theta_A(n_A(h), n_A(l_{A_p})) \mid n_B(l_{B_p}) \supset n_B(h) \} \leq (1/k) \lambda_{B_{p+1}}$$

et par (1.4) et (3.2):

$$\sin^2 \theta_A(n_A(l_{B_p}), n_A(l_{A_p})) \leq K \lambda_{A_{p+1}} / (k \lambda_{A_p}) \quad (3.5)$$

c.q.f.d., par (1.5).

La relation (3.5) exprime la relation entre sous-espaces de contraintes définis suivant les deux métriques A et B.

4 Conclusions

Ces résultats, en un sens, justifient l'usage répandu de ne pas chercher à préciser outre mesure la métrique utilisée. Le rapport K/k caractérise par (3.1) les variations admissibles sur la métrique. Les résultats ci-dessus montrent que la hiérarchie des sous-espaces de composantes principales n'est stable par changement de métrique que s'il existe dans l'échelonnement des valeurs propres λ de l'inertie une séparation correspondant à un rapport nettement supérieur à K/k .

Il existe dans certains cas plusieurs séparations de ce type, donc plusieurs sous-espaces stables de composantes principales. On peut avoir, par exemple, un sous-espace stable de grandes composantes, et un autre sous-espace stable, englobant le premier, de composantes grandes ou moyennes. Il serait vain de vouloir chercher dans ce dernier un complément stable du premier, contenant uniquement des composantes moyennes : il est facile de construire des contre-exemples montrant l'instabilité du résultat.

Enfin, notons qu'il est facile de passer, par inversion des métriques, aux sous-espaces stables du dual de E, ce qui permet de réexprimer les résultats ci-dessus en termes d'ensembles plus ou moins sévères de contraintes sur le nuage étudié. Cette formulation peut être plus commode que celle des composantes principales dans des problèmes tels que la reconnaissance des formes.