ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Mathématiques

GUY FOURT

Étude géométrique de l'analyse de la variance multidimensionnelle

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 67, série *Mathématiques*, n° 17 (1979), p. 44-62

<http://www.numdam.org/item?id=ASCFM_1979__67_17_44_0>

ETUDE GEOMETRIQUE DE L'ANALYSE DE LA VARIANCE MULTIDIMENSIONNELLE

Guy FOURT

Université de Clermont-Ferrand

SOMMAIRE : Nous présentons une construction géométrique du test de Pillai et certaines généralisations.

INTRODUCTION :

Cette étude prolonge le travail [4], publié dans ces mêmes pages. Rappelons qu'il s'agit, étant données des v.a. $X_i (i = 1,2,...j)$ indépendantes à valeurs dans R^k, de lois $\mathcal{N}(m_i, \Lambda)$, et une matrice de Wishart S de paramètre Λ à ρ-j degrés de liberté, $(\rho - j > k)$ de construire des tests de l'hypothèse $\{m_i = 0 \ \forall i\}$ (notée H_o) contre $\{\exists i : m_i = 0\}$ (notée $\overline{H}_o$). Sans perte de généralité on peut supposer Λ inversible ; de plus on se restreindra aux tests invariants par l'action des transformations linéaires inversibles $X_i \rightsquigarrow G(X_i)$ et $S \rightsquigarrow G\ S\ {}^tG$. En conclusion de [4], nous laissions entendre que cela risquait de conduire à des tests sophistiqués ; de fait nous retrouverons, sous une forme cachée, le test de Pillai. Signalons qu'on trouve dans un article de R. SCHWARTZ [7], repris sous une forme plus générale par TAKEAKI KARIYA [8] une variante de ce travail, non bayesienne, mais utilisant l'invariance sous un groupe de transformations plus riche. Ces deux types de résultats expliquent le travail de PILLAI et JAYACHANDRAN [6].

Les notations utilisées ici seront les mêmes qu'en [4], et nous ne reviendrons pas sur la description des cônes homogènes qui a été faite ; le chapitre V montrera le rôle fondamental de cette structure en abordant des généralisations, les quatre premiers étant l'étude détaillée du problème de l'analyse de la variance.

I - RECHERCHE DE STATISTIQUES INVARIANTES MAXIMALES

Soit T une transformation triangulaire. Sous son action, les données sont transformée en $T(X_i)$ et $T\ S^t T$. Les statistiques invariantes sous l'action du groupe triangulaire sont donc les $S^{\perp -1}\perp(X_i)$. Soit maintenant pour G élément du groupe orthogonal O(k), T et T' les matrices triangulaires supérieures définies par $T^tT = S$, $T'^{t}T' = G\ S^t\ G = S'$.
Si les X_i sont transformés en $G(X_i)$ alors $S^{\perp -1}\perp(X_i) = T^{-1}(X_i)$ est transformé en $S'^{\perp -1}\perp(G(X_i)) = T'^{-1}\ G(X_i)$ ou encore en $T'^{-1}G\ T(S^{\perp -1}\perp(X_i))$, et l'opération $T'^{-1}G\ T$ est une transformation orthogonale. Rechercher un invariant par l'ensemble de ces opérations revient à rechercher un invariant par le groupe engendré par ces opérations (elles ne forment pas un groupe !).

THEOREME 1 :

Le groupe engendré par les $T'^{-1}G\ T$ est O(k) .

Preuve : Montrons que pour T fixé, $G \rightsquigarrow T'^{-1}G\ T$ est une bijection entre deux voisinages de l'identité, ou ce qui est équivalent, que $d\ G \rightsquigarrow d(T'^{-1}\ G\ T)$ est une bijection entre algèbres de Lie de O(k).
Les calculs seront effectués en G = identité, d'où T' = T.

$$d(T'^{-1}\ G\ T) = d(T'^{-1})\ T + T^{-1}\ dG\ T = -\ T'^{-1}\ dT' + T^{-1}\ dG\ T.$$

Différentions $T'^{t}T' = G\ T^t\ T^t\ G$

$$d\ T'^{t}\ T + T\ d\ T' = d\ G\ T^t\ T + T^t\ T\ d\ G.$$

Les matrices $d\ T'^{t}T$ et $d\ G\ T^t\ T$ ont donc même symétrisées et diffèrent d'une matrice antisymétrique (donc dans l'algèbre de Lie de O(k)) dh

$$d\ T'^{t}T = d\ G\ T^tT + dh.$$

On en déduit

$d(T'^{-1}G\ T) = T^{-1}\ dh^t\ T^{-1}$, et le fait que $dh \rightsquigarrow d(T'^{-1}\ GT)$ est une bijection linéaire. De plus $T^{-1}d\ T' = T^{-1}\ d\ G\ T + T^{-1}\ dh^t\ T^{-1}$ doit être triangulaire supérieure. Si $(\ )^-$ désigne la partie triangulaire strictement inférieure à la diagonale d'une matrice, dh vérifie donc :

$$(T^{-1}\ d\ G\ T)^- + (T^{-1}\ dh^t\ T^{-1})^- = 0.$$

La correspondance $dh \rightsquigarrow dG$ est linéaire, et si $dG = O\ T^{-1}\ dh^t\ T^{-1}$ qui est antisymétrique ne peut être que nulle, de même que dh ; $dh \rightsquigarrow dG$ est donc une bijection.

La figure formée par O et les $S^{\perp -1}{}_{\perp}(X_i)$ est donc globalement invariante. Nous étudierons donc dans un premier temps, les $||S^{\perp -1}{}_{\perp}(X_i)||$ et les angles entre les $S^{\perp -1}\ {}_{\perp}(X_i)$.

Notons maintenant par $H_{(r,r+1)}(\theta)$ la matrice de rotation d'angle θ conservant toutes les coordonnées sauf celles d'ordre r et r+1 et de façon générale définissons par récurrence :

$$H_{(r,r')}(\theta_r, \theta_{r+1},\dots, \theta_{r'-1}) = H_{(r,r'-1)}(\theta_r,\dots, \theta_{r'-2})\ H_{(r'-1,r')}(\theta_{r'-1}).$$

Citons les deux faits suivants :

a) On peut établir par récurrence la décomposition d'Euler de toute matrice H de O(k) :

$$H = H_{(1,k-1)}\ (\theta_1,\dots, \theta_{k-2}) \cdot H_{(k-1,k)}\ (\theta_k) \cdot H'$$

où $H' \in O(k-1)$ (décomposition precession x nutation x rotation propre), d'où la décomposition :

$$H = H_{(1,k)}\ H_{(1,k-1)} \cdots\ H_{(1,i)}\ \cdots\ H_{(1,2)}.$$

b) En explicitant les arguments $\theta_{r,i}$ $(r < i)$ de $H_{(1,i)}$ dans la décomposition ci-dessus, la mesure de Haar dH de O(k) est donnée par :

$$dH = \prod_{0<r<i} |\sin \theta_{r,i}|^{r-1} \prod_{0<r<i} d\theta_{r,i}\ .$$

La technique utilisée consistera de façon naturelle à passer dans un repère aléatoire mettant en évidence des invariants équivalents comme suit :
Aux $S^{\perp -1}{}_{\perp}(X_i)$ on applique le procédé d'orthogonalisation de Schmidt ; si $j < k$, on complète par un repère choisi "au hasard" dans l' orthogonal de ces vecteurs. (Pour des raisons de commodité de calcul on réordonnera à l'envers ce repère). On écrit alors :

$$S^{\perp -1} \perp(X_1) = ||S^{\perp -1} \perp(X_1)|| \, v_k,$$

et plus généralement :

$$S^{\perp -1}\perp(X_i) = ||S^{\perp -1}\perp(X_i)|| H_{(k-i+1\vee 1,k)} \, (\theta^{(i)}_{k-i+1\vee 1},\ldots,\theta^{(i)}_{k-1}) \quad [v_k].$$

Les angles $\theta_r^{(i)}$ sont encore des invariants (par construction) et leur donnée permet de calculer les angles entre les $S^{\perp -1}\perp(X_i)$ dans le repère aléatoire, donc dans tout repère orthonormé. Ce sont eux que nous utiliserons de préférence.

II - CALCUL DE LA LOI DES STATISTIQUES INVARIANTES

THEOREME 2 :

La densité des $T_i = ||S^{\perp -1}(X_i)||^2$ et des $\theta_r^{(i)}$ est donnée par :

$$k(e)e^{-\sum\limits_{i=1}^{j}||a\perp(m_i)||^2} \prod_{i=1}^{j} t_i^{\frac{k}{2}-1} \prod_{i=1}^{j} \prod_{r=k-i+1\vee 1}^{k-1} |\sin\theta_r^{(i)}|^{r-1}$$

$$\iint_{\mathcal{M}_k^+ \times O(k)} e^{-<e,s\perp(e+\sum\limits_{i=1}^{j} t_i H \,.\, H_i(e_k) \Delta H \,.\, H_i(e_h)>}$$

$$e^{2\sum\limits_{i=1}^{j} \sqrt{t_i} <a\perp(m_i),\, s\perp H \,.\, H_i(e_k)>} \; s^{\frac{\rho}{2}(1\ldots 1)} \; d\mu(s)\, dH$$

où on a abrégé $H_{(k-i+1\vee 1,k)}$ en H_i et où $a = \dfrac{\Lambda^{-1}}{4}$.

Preuve : $P(\bigcap\limits_{i=1}^{j} \{T_i < t_i\} \cap \bigcap\limits_{r=h-i+1\vee 1}^{k-1} \{\theta_r^{(i)} \in A_r^i\})$

$$= k(a)\, e^{-\sum\limits_{i=1}^{j}||a\perp(m_i)||^2} \iint_{||s^{\perp -1}\perp(x_i)||^2 < t_i,\, \theta_r^{(i)} \in A_r^i} e^{-<a^*,\, s\perp(e+\sum\limits_{i=1}^{j} x_i \Delta x_i)>}$$

$$e^{2<a^*,\, \sum\limits_{i=1}^{j} x_i \Delta m_i>} \; s^{\frac{\rho-j}{2}(1,\ldots,1)} \; d\mu(s) \prod_{i=1}^{j} dx_i .$$

On transforme cette expression, comme dans [4] pour obtenir, $d\sigma$ étant la mesure de Lebesgue sur $||u|| = 1$:

$$h(e)\, e^{-\sum_{i=1}^{j} ||a\perp(m_i)||^2} \iint_{\tau_i < t_i\ ,\ \theta_r^{(i)} \in A_r^i} e^{-\langle e,\ s\perp(e + \sum_{i=1}^{j} \tau_i\, u_i\, \Delta u_i)\rangle}$$

$$e^{2 \sum_{i=1}^{j} \sqrt{\tau_i}\, \langle a\perp(m_i),\ s\perp(u_i)\rangle}\ {}_s\frac{\rho}{2}\,(1\ldots 1)\ \prod_{i=1}^{j} \tau_i^{\frac{k}{2}-1}$$

$$d\mu(s) \prod_{i=1}^{j} d\,\tau_i \prod_{i=1}^{j} d\sigma(u_i)\ .$$

Associons, comme en I aux u_i, un repère orthogonal tel que $u_1 = v_k$, u_2 soit combinaison linéaire de $(v_k,\ v_{k-1})\ldots$; soit H la matrice de changement de base et H_i la matrice telle que :

$$H_{(k-i+1V1,k)}[vk] = u_i.$$

Le résultat est alors obtenu par le lemme suivant :

<u>Lemme du changement de repère aléatoire</u> :

Avec les notations précédentes :

$$\prod_{i=1}^{j} d\sigma\ (u_i) = dH \prod_{i=1}^{j} \prod_{r=k-i+1V1}^{k-1} |\sin\theta_r^{(i)}|^{r-1}\ d\theta_r^{(i)}\ .$$

<u>Preuve du lemme</u> : Si les u_i suivent la loi $\prod_{i=1}^{j} d\sigma(u_i)$, il est évident que la loi de H, conservée par translation à gauche est dH. Ecrivons la décomposition d'Euler : $H = H^*_{(1,k)} \cdot h_1$ $(h_1 \in O(k-1))$.

Comme $h_1(e_k) = e_k$ et que $u_1 \rightsquigarrow H_{(1k)}(e_k)$ est bijective, la loi de h_1 conditionnelle à u_1 est la loi de h_1 conditionnelle à $H^*_{(1,k)}$ donc la mesure de Haar de O(k-1). On écrit alors :

$$u_2 = H^*_{(1k)} \cdot h_1 \cdot H_{(k-1,k)}\ (\theta^{(2)}_{k-1})\quad [e_k]\ .$$

La loi de u_2 conditionnelle à u_1 est uniforme ; donc la loi de

$h_1 \cdot H_{(k-1,k)}\ (\theta^{(2)}_{k-1})\ \ [e_k]$ doit être uniforme. Puisque $h_1(e_{k-1})$ est uniforme sur la sphère du plan orthogonal à e_k, $\theta^{(2)}_{k-1}$ suit dont la loi :

$$|\sin \theta^{(2)}_{k-1}|^{\ k-2} \quad d\theta^{(2)}_{k-1}.$$

Plus généralement, par décompositions d'Euler successives, nous écrivons :

$$H = H^*_{(1,k)}\ H^*_{(1,k-1)} \ldots . \ \cdot H^*_{(1,k-i+2\vee 2)} \cdot h_{i-1} \ .$$

La loi de h_{i-1}, conditionnelle aux $(u_1,\ldots u_{i-1})$ ou encore aux

$(H^*_{(1,k)}\ H^*_{(1,k-1)} \ldots\ H^*_{(1,k-i+2\vee 2)})$ est la mesure de Haar de $O(k-i+1\vee 1)$ et

on en déduit de ce que la loi de u_i conditionnelle aux $(u_1,\ldots\ u_{i-1})$ est uniforme et de ce que,

$$h_{i-1} \cdot H_{(k-i+1\vee 1,k)}\ (\theta^{(i)}_{k-i+1\vee 1},\ \ldots\ \theta^{(i)}_{k-1}) \qquad [e_k]$$

doit être uniforme, que la loi de $\theta^{(i)}_r$ est :

$$\prod_{r=k-i+1}^{k-1} \quad |\sin \theta^{(i)}_r|^{r-1} \quad d\theta^{(i)}_r .$$

<u>Corollaire</u> :

Sous l'hypothèse nulle la valeur de l'intégrale écrite au théorème 2 est

$$f(t_1\ldots t_j) = (e + \sum_{i=1}^{j} \ t_i\ H_i(e_k)\ \Delta\ H_i(e_k))^{-\frac{\rho}{2}\,(1\ldots 1)} \ .$$

<u>Preuve</u> : Posons $w = e + \sum_{i=1}^{j} t_i\ H_i(e_k)\ \Delta\ H_i(e_k)$, nous avons à calculer, pour $w \in \mathcal{M}^+_k$:

$$\iint_{\mathcal{M}^+_k \times O(k)} e^{-\langle e, s\perp(H-w^tH)\rangle}\ s^{\frac{\rho}{2}(1\ldots 1)} \quad d\mu(s)\ dH$$

qui se réduit facilement (voir [2] ou [3] à

$$\Gamma(\frac{\rho}{2}\ (1\ldots 1)) \int_{O(h)} \frac{dH}{(H\ w^tH)^{\alpha - \frac{\rho}{2}(1\ldots 1)}}) \ ,$$

Notons par $\det_m(s)$ le déterminant de la matrice $m \times m$ inférieure droite de s ; en explicitant α, (voir [3]) l'intégrale se met sous la forme :

$$\frac{1}{(\det_k w)^{\frac{\rho-k+1}{2}}} \int_{O(k)} \frac{dH}{\prod_{m=1}^{k-1} \det_m (Hw\,{}^tH)} \; .$$

Une récurrence sur k va montrer que la dernière intégrale est proportionnelle à $(\det_k w)^{-\frac{k-1}{2}}$.

Sans perte de généralité, on peut supposer que w est diagonale de valeurs propres $a_1, \ldots , a_k$; H admet une décomposition (de type Euler)

$$H = \overline{H} \circ H_{(1,k)} \quad \text{où} \quad \overline{H}\,(e_1) = e_1 .$$

Nous avons à calculer d'abord, en posant $w_1 = H_{(1,k)}\; w\; {}^t H_{(1,k)}$

$$\int \frac{d\,\overline{H}}{\prod_{m=1}^{k-1} \det_m (\overline{H}\, w_1\, {}^t\overline{H})} \quad \text{qui, puisque}$$

$$\det_{k-1} (\overline{H}\, w_1\, {}^t\overline{H}) = \det_{k-1} w_1$$

et par l'hypothèse de récurrence, est égale à $\dfrac{1}{(\det_{k-1} w_1)^{\frac{k}{2}}}$.

Posons ensuite $w_2 = H_{(2,k)}\; w\,{}^tH_{(2,k)}$, qui s'écrit encore, en isolant les deux premières lignes et colonnes, sous la forme "blocs"

$$\begin{pmatrix} a_1 & 0 & 0 \\ 0 & \alpha_2 & {}^t w_3' \\ 0 & w_3' & w_3 \end{pmatrix}$$

On en déduit que :

$$\det_{k-1} H_{(1,2)}(\theta).\; w_2\,{}^tH_{(1,2)}(\theta) = \begin{vmatrix} a_1 \sin^2\theta + \alpha_2 \cos^2\theta, & \cos\theta\,{}^t w_3' \\ \cos\theta\; w_3' & w_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 \sin^2\theta & \cos\theta\, {}^t w_3' \\ 0 & w_3 \end{vmatrix} + \begin{vmatrix} \alpha_2 \cos^2\theta & \cos\theta^t\, w_3' \\ \cos\theta\, w_3' & w_3 \end{vmatrix}$$

$$= a_1 \sin^2\theta \ \det_{k-2} w_2 \ + \ \cos^2\theta \det_{k-1} w_2.$$

On calcule alors $\displaystyle\int \frac{|\sin\theta|^{k-2} \quad d\theta}{[\det_{k-1} [H_{12}(\theta).\ w_2\ {}^tH_{12}(\theta)]]^{\frac{k}{2}}}$

par des méthodes élémentaires pour trouver (à un facteur près)

$$a_1^{-\frac{k-1}{2}} (\det_{k-1} w_2)^{-\frac{1}{2}} \quad (\det_{k-2} w_2)^{-\frac{k-1}{2}}$$

soit encore $a_1^{-\frac{k-1}{2}} (a_2 \ldots a_k)^{-\frac{1}{2}} (\det_{k-2} w_2)^{-\frac{k-1}{2}}$.

les intégrales à calculer ensuite sont du même type, en remplaçant k par k-1..., k-2,... d'où le résultat :

$$a_1^{-\frac{k-1}{2}} (a_2 \ldots . a_k)^{-\frac{1}{2}} \quad a_2^{-\frac{k-2}{2}} (a_3 \ldots a_k)^{-\frac{1}{2}} \ldots\ldots a_k^{-\frac{1}{2}}$$

$$= a_1^{-\frac{k-1}{2}} a_2^{-\frac{k-1}{2}} \ldots\ldots a_k^{-\frac{k-1}{2}} = (\det w)^{-\frac{k-1}{2}} .$$

III - RECHERCHE DE TESTS LOCALEMENT UMP

THEOREME 3 :

Il n'existe pas de tests localement UMP invariants sous le groupe linéaire de H_o contre $\overline{H}_o$.

Preuve : Nous différencions la densité de la statistique invariante, par rapport au paramètre :

$$m = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_j \end{pmatrix}$$

en m = 0. La différentielle première est identiquement nulle.

L'expression $f(t_1, \dots, t_j)$ étant définie comme précédemment, nous posons :

$$M_{ii'} = \iint_{\mathscr{M}_k^+ \times O(k)} e^{-\langle e, s\perp(e + \sum_{i=1}^{j} t_i\, H.H_i(e_k)\, \Delta\, H.H_i(e_k))\rangle}$$

$$\sqrt{t_i t_{i'}}\, s\perp(H.H_i(e_k)\ \Delta\ H.H_{i'}(e_k))\, s^{\frac{\rho}{2}(1\dots1)}\, d\mu(s)\, dH.$$

La matrice de la différentielle seconde se calcule par blocs k x k au moyen des différentielles partielles $D^2_{m_i\, m_{i'}}$. La forme différentielle seconde se met sous la forme :

$$(t_{\mu_1} \dots t_{\mu_j})\, D^2_{(m,m)} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_j \end{pmatrix} = \sum_{1 \le i \le i' \le j} {}^t(a\perp(\mu_{i'})\, A_{ii'}(a\perp(\mu_i))$$

$$= \sum_{i=1}^{j} {}^t(a\perp(\mu_i)\, A_{ii}(a\perp(\mu_i)) + \sum_{1 \le i < i' < j} {}^t(a\perp(\mu_{i'})\, (A_{ii'} + A_{i'i})\, (a\perp(\mu_i))$$

avec $A_{ii} = 2\, f(t_1 \dots t_j)\, e + 4\, M_{ii}$

$A_{ii'} + A_{i'i} = 4\, M_{ii'}$ pour $i \neq i'$.

Etant donné un élément $\overline{H}$ dans O(k), on effectue le changement de variable

$$\begin{cases} s' = \overline{H}.\ s.\,{}^t\overline{H} \\ H' = T'^{-1}.\ \overline{H}.\ T.\ H \end{cases}$$

(où on a posé $s = T.{}^tT$, $s' = T'.{}^tT'$, T et T' étant des matrices triangulaires supérieures). $s \rightsquigarrow \overline{H}.\ s.\,{}^t\overline{H}$ conserve $s^{\frac{\rho}{2}(1\dots1)}\, d\mu(s)$, et pour s fixé $T'^{-1}.H.T.$ est orthogonale et la deuxième expression conserve dH ; la première expression ne faisant pas intervenir H, l'ensemble des deux conserve $s^{\frac{\rho}{2}(1\dots1)}\, d\mu(s)dH$. On en déduit les égalités $\overline{H}.M_{ii'}\,{}^t\overline{H} = M_{ii'}$ qui montrent que les $M_{ii'}$ sont proportionnelles à la matrice identité e. soit :

$$M_{ii'} = \frac{\langle e, M_{ii'}\rangle}{k}\, e$$

Pour $i \neq i'$ les produits scalaires $<e, M_{ii'}>$ sont non nuls (ils sont évidemment strictement positifs en $H_i = H_{i'}$, donc strictement positifs dans un voisinage de cette variété). En appliquant le lemme de Neymann Pearson, un test localement M.P. sera fortement influencé par le choix des produits scalaires $<a\perp(m_{i'})\ a\perp(m_i)>$ et il n'existe pas de test localement U.M.P.

<u>THEORME 4</u> :

Si les μ_i sont des v.a. indépendantes, de même loi, et si leur loi commune ne dépend que de $||a\perp(\mu_i)||$ il existe un test localement UMP de H_0 contre les hypothèses $||a\perp(\mu_1)|| = \dots = ||a\perp(\mu_j)||$.

Sa zone de rejet est donnée par :

$$- \sum_{i=1}^{j} t_i \frac{\partial f}{\partial t_i} > c\ f.$$

<u>Preuve</u> : Les hypothèses sur les μ_i impliquent que :

$$E(a\perp(\mu_i))\ |\ ||a\perp(\mu_1)||\ \dots\ ||a\perp(\mu_j)||) = 0$$

$$E(||a\perp(\mu_i)||^2\ |\ ||a\perp(\mu_1)||\dots\dots ||a\perp(\mu_j)||) = ||a\perp(\mu_i)||^2.$$

Le lemme de Neymann Pearson conduit à maximiser :

$$E(D^2_{mm}[\mu,\mu]\ |\ ||a\perp(\mu_1)||\dots ||a\perp(\mu_j)||)$$
$$= \sum_{i=1}^{j} ||a\perp(\mu_i)||^2 \left(2 f - 4 \frac{t_i}{k} \frac{\partial f}{\partial t_i}\right)$$

et à rejeter H_0 si cette expression est supérieure à c_1 f. Supposant les $||a\perp(\mu_i)||^2$ égaux entre eux la zone de rejet peut s'écrire sous la forme annoncée.

<u>THEOREME 5</u> :

Sous les mêmes hypothèses Bayesiennes, il n'existe pas de test UMP invariant. (pour $j > 1$).

<u>Preuve</u> : Choisissons pour loi des m_i la loi $\mathcal{N}(0, \varepsilon\Lambda)$ $(\varepsilon > 0)$. Par les formules de convolution, la loi moyenne de X_i sous l'hypothèse adverse sera $\mathcal{N}(0, (1+\varepsilon)\Lambda)$.

Reprenant les calculs précédents, la densité de la statistique invariante est :

$$\frac{K \prod_{i=1}^{j} t_i^{\frac{k}{2}-1} \prod_{i=1}^{j} \prod_{r=k-i+1V1}^{k-1} |\sin \theta_r^{(i)}|^{r-1}}{(\det(e+ \frac{1}{1+\varepsilon} \sum_{i=1}^{j} t_i H_i(e_k) \Delta H_i(e_k)))^{\frac{\rho}{2}}} .$$

Le test M.P. correspondant sera basé sur le rapport :

$$\frac{\det (e + \sum_{i=1}^{j} t_i V_i \Delta V_i)}{\det (e(1+\varepsilon) + \sum_{i=1}^{j} t_i V_i \Delta V_i))}$$

et dépend de ε pour $j > 1$.

IV - OŬ ON RETROUVE LE TEST DE PILLAI

THEOREME 6 :

Le test construit au théorème 4 n'est autre que le test de la trace de Pillai, donné par :

$$\sum_{i=1}^{j} {}^t X_i \, (S + \sum_{i=1}^{j} X_i \Delta X_i)^{-1} X_i > C.$$

Preuve : Nous avons établi que (corollaire du théorème 2)

$$f = K \det (e + \sum_{i=1}^{j} t_i V_i \Delta V_i)^{-\frac{\rho}{2}} .$$

La zone de rejet peut donc s'écrire :

$$\sum_{i=1}^{j} t_i \frac{\partial}{\partial t_i} (\det (e + \sum_{i=1}^{j} t_i V_i \Delta V_i)) > C_2 \det (e + \sum_{i=1}^{j} t_i V_i \Delta V_i).$$

Notons par $\underline{s}_m$ le $m^{\text{ième}}$ vecteur colonne de la matrice s. On a alors les égalités :

$$\sum_{i=1}^{j} t_i \frac{\partial}{\partial t_i} (\det (e + \sum_{i=1}^{j} t_i V_i \Delta V_i))$$

$$= \sum_{i=1}^{j} t_i \sum_{m=1}^{k} \det \left(\underline{e+ \sum_{i=1}^{j} t_i V_i \Delta V_i}_{\,1}, \ldots, \underline{e+ \sum_{i=1}^{j} t_i V_i \Delta V_i}_{\,m-1}, \underline{V_i \Delta V_i}_{\,m}, \right.$$
$$\left. , \underline{e+ \sum_{i=1}^{j} t_i V_i \Delta V_i}_{\,m+1}, \ldots \underline{e+\sum_{i=1}^{j} t_i V_i \Delta V_i}_{\,k} \right)$$

$$= \sum_{m=1}^{k} \det\left(\frac{e+\sum_{i=1}^{j} t_i \; V_i \Delta V_i}{} 1, \ldots\ldots, \frac{e+\sum_{i=1}^{j} t_i \; V_i \Delta V_i}{} m-1, \; \sum_{i=1}^{j} \frac{t_i V_i \Delta V_i}{} m, \right.$$

$$\left. \frac{e+\sum_{i=1}^{j} t_i \; V_i \Delta V_i}{} m+1, \; \ldots\ldots, \; \frac{e+\sum_{i=1}^{j} t_i \; V_i \Delta V_i}{} k \right)$$

$$= \sum_{m=1}^{k} \det\left(e - \frac{e_m \Delta e_m}{2} + \sum_{i=1}^{j} t_i \; V_i \Delta V_i \right).$$

Comme det est une fonction puissance, la zone de rejet peut alors s'écrire :

$$\sum_{m=1}^{k} \det \left(e - \left(e + \sum_{i=1}^{j} t_i \; V_i \Delta V_i\right)^{\perp -1} \perp \frac{e_m \Delta e_m}{2} \right) > C_2.$$

On utilise alors l'identité, où $u \in R^k$:

$\det(e - u \Delta u) = 1 - ||u||^2 = 1 - \langle e, u \Delta u\rangle$

pour ramener la zone de rejet à :

$$k - \sum_{m=1}^{k} \left\langle e, \left(e+ \sum_{i=1}^{j} t_i \; V_i \Delta V_i\right)^{\perp -1} \perp \frac{e_m \Delta e_m}{2} \right\rangle > C_2$$

ou plus simplement à :

$$\left\langle e, e - \left(e + \sum_{i=1}^{j} t_i \; V_i \Delta V_i\right)^{\perp -1} \right\rangle > C_2.$$

Remarquant alors que $(e + s)^{\perp -1} \perp s = (e+s)^{\perp -1} \perp (e+s-e) = e - (e+s)^{\perp -1}$

on écrit la zone de rejet comme :

$$\left\langle e, \left(e+ \sum_{i=1}^{j} t_i \; V_i \Delta V_i\right)^{\perp -1} \perp \left(\sum_{i=1}^{j} t_i \; V_i \Delta V_i\right)\right\rangle > C_2$$

et rappelant que $t_i \; V_i \Delta V_i = S^{\perp -1} \perp (X_i \Delta X_i)$ comme :

$$\left\langle e, \left(S + \sum_{i=1}^{j} X_i \Delta X_i\right)^{\perp -1} \perp \sum_{i=1}^{j} X_i \Delta X_i \right\rangle > C_2$$

qui est la forme classique, compte tenu de la définition de $\langle \; , \; \rangle$ en terme de trace.

V - EXTENSION A CERTAINES HYPOTHESES LINEAIRES

A quelques détails près (essentiellement de mise en forme), ces résultats peuvent s'appliquer à une classe de cônes homogènes plus vaste. De façon précise, nous supposerons que le paramètre Λ^{-1} appartienne à un sous espace vectoriel $\mathcal{P}$ de $R^{\frac{k(k+1)}{2}}$, et surtout que le cône $U^*=\mathcal{P}\cap\mathcal{M}_k^+$ est homogène sous l'action de certaines transformations linéaires,

$$a^*\in\mathcal{P}\cap\mathcal{M}_k^+ \rightsquigarrow {}^t g \;\; a^* \; g.$$

Nous aurons fréquemment à faire référence au cône homogène $\mathcal{M}_k^+$; nous repérerons par un indice 0 les outils de cet espace. Π désignera la projection de $R^{\frac{k(k+1)}{2}}$ sur $\mathcal{P}$.

On pourra se reporter à [3] pour l'étude systématique de ces outils.

LEMME 1 :

Sans perte de généralité, on peut supposer que la loi de groupe $\perp$ sur U^* est la restriction aux éléments de $\mathcal{P}$ de la loi de groupe $\perp_0$ sur $\mathcal{M}_k^+$ (en particulier que $e \in U^*$).

Preuve : Soit $\mathcal{G}$ un sous groupe de transformations $a^*\rightsquigarrow {}^t g \; a^* g$ agissant de façon simplement transitive. $\mathcal{G}$ est résoluble et il en est de même du groupe des transformations linéaires associées : elles sont donc simultanéments triangulables et s'écrivent sous la forme $g = \bar{g}\, T\, \bar{g}^{-1}$. Un changement de base de matrice $\bar{g}$ ramène au cas d'un sous groupe d'opérateurs triangulaires.

LEMME 2 :

La loi Δ^* sur $\mathcal{P}$ est la restriction de la loi Δ_0^* sur $R^{\frac{k(k+1)}{2}}$. La loi duale Δ sur $\mathcal{P}$ est définie par $a \,\Delta\, b = \Pi(a \;\Delta_0\; b)$. Enfin les opérateurs $u \rightsquigarrow \Pi(g\; u\; {}^t g)$ forment le groupe agissent de façon simplement transitive sur le cône dual $\mathcal{U}$ de $\mathcal{U}^*$.

Preuve : La première assertion se déduit sans peine du lemme 1. L'opérateur $L^*_{o\begin{pmatrix}a\\o\end{pmatrix}}$, agissant sur $R^{\frac{k(k+1)}{2}} = \mathscr{P} \otimes \mathscr{P}^{\perp}$ s'écrit donc $\begin{pmatrix} L^*_a & M_1 \\ 0 & M_2 \end{pmatrix}$.

Son dual relativement au produit scalaire $< , >_o$ sur $R^{\frac{k(k+1)}{2}}$ est donc $\begin{pmatrix} L_a & 0 \\ M_1^* & M_2^* \end{pmatrix}$ et on vérifie ensuite que $\begin{pmatrix} L_a(b) \\ 0 \end{pmatrix} = \Pi\, L_{o\begin{pmatrix}a\\o\end{pmatrix}}\left[\begin{pmatrix}b\\0\end{pmatrix}\right]$.

Enfin $\Pi\, L_{o\begin{pmatrix}a\\o\end{pmatrix}}\, \Pi = \Pi\, L_{o\begin{pmatrix}a\\o\end{pmatrix}}$. On en déduit que :

$$(\Pi\, L_{o\begin{pmatrix}a\\o\end{pmatrix}})^n = \Pi(L_{o\begin{pmatrix}a\\o\end{pmatrix}})^n,$$ et par passage à l'exponentielle que

$\exp(\Pi\, L_{o\begin{pmatrix}a\\o\end{pmatrix}}) = \Pi \exp L_{o\begin{pmatrix}a\\o\end{pmatrix}}$, d'où la dernière assertion puisque les $\exp L_a$ engendrent le groupe simplement transitif de $\mathscr{U}$.

LEMME 3 :

Les restrictions des fonctions puissances de $\mathscr{M}_k^+$ à $\mathscr{U}^*$ sont les fonctions puissances de $\mathscr{U}^*$. (Conséquence facile du lemme 1).

DEFINITION :

Pour u et v dans R^k, on définit $u \Delta v = \Pi(u \Delta_o v)$ (c'est un élément de $\mathscr{P}$), et si s est un élément de $\mathscr{U}$ se mettant sous la forme (unique) $s = \Pi(T^tT)$, on posera $s\perp(u) = T(u)$.

THEOREME 7 :

Pour un échantillon de v.a. normales $\mathscr{N}(m,\Lambda)$, et une matrice de Wishart S_o à $\rho-1$ degrés de liberté de paramètre Λ, les statistiques $X = \frac{\overline{X}}{\sqrt{2}}$ et $S = \Pi(S_o)$ sont exhaustives si $\Lambda^{-1} \in \mathscr{P}$; leurs densités prennent alors la forme (avec $a^* = \frac{\Lambda^{-1}}{4}$)

$$\frac{1}{\sqrt{2n}\, a^{\beta}} \quad e^{-<a^*,(x-m)\,\Delta\,(x-m)>} \quad dx \qquad (1)$$

$$\frac{1}{\Gamma((\rho-1)\;)\;a^{(\rho-1)\beta}} e^{-<a^*,\;s>} \; s^{(\rho-1)} \; d\mu(s) \qquad (2)$$

Preuve : La première assertion est classique (voir exemple [1] chapitre X).
Par le lemme 3, det a^* est (pour les éléments de $\mathcal{U}^*$) une fonction puissance notée ici $a^{*\,s(2\beta)} = a^{2\beta}$, de plus :

$$<a^*,\;(x-m)\;\;\Delta_o(x-m)>_o \;=\; <a^*,\Pi[(x-m)\;\Delta_o(x-m)]\;>_o$$

$$= \;<a^*,(x-m)\;\Delta\;(x-m)> \quad ,$$

d'où la forme de la loi (1).

S_o suit la loi de Wishart à $(\rho-1)$ degrés de liberté de transformée de Laplace $\left[\det(\;a^* + p\;)\right]^{-\frac{\rho-1}{2}}$.

La transformée de Laplace de $\Pi(S_o)$ est donc, pour p dans $\mathcal{U}^*$:

$$E(e^{-<p,\Pi(S_o)>} \;=\; E(e^{-<p,S_o>_o}$$

$$= (\det(a^* + p))^{-\frac{\rho-1}{2}} \;=\; (a^* + p)^{(\rho-1)\;s(\beta)}$$

et le résultat (2) s'obtient alors par identification des transformées de Laplace. cf[2] ou [3] .

Nous allons maintenant suivre pas à pas le début de cet exposé, quitte à l'aménager un peu pour faire intervenir le caractère propre de $\mathcal{U}$. Aussi les théorèmes seront énoncés sans autre justification que les corrections introduites.

$X_i\,(i=1,2\ldots j)$ désignera des v.a. indépendantes de R^k de loi (1) et S une v.a. de loi (2) indépendantes des X_i $\mathcal{K}$ sera le sous groupe de $O(k)$ conservant $\mathcal{P}$.

THEOREME 8 :

Une statistique invariante maximale sous le groupe linéaire conservant $\mathcal{P}$ est donnée par les $||S^{\perp-1}(X_i)||^2$, les $H^{(i)}_{(k-j\vee 1,k)}$ et par $H(\mathrm{mod}\;\mathcal{K})$.

Preuve : Les $S^{\perp-1}(X_i)$ sont les seules statistiques invariantes sous les transformations triangulaires conservant $\mathcal{P}$, par construction de $\perp$. Sous l'action d'une transformation K de $\mathcal{K}$, elles subissent l'action d'une autre transformation K_1 de $\mathcal{K}$. La matrice de passage aléatoire H devient alors K_1 H. Le résultat s'obtient en remarquant que les K_1 engendrent $\mathcal{K}$ (théorème 1).

THEOREME 9 :

Ces statistiques ont pour densité, relativement aux d t_i, $d\theta_r^{(i)}$ et, à la mesure relativement invariante sur $O(k)/\mathcal{K}$

$$k(e)\, e^{-\sum_{i=1}^{j} ||a\perp(m_i)||^2} \prod_{i=1}^{j} t_i^{\frac{k}{2}-1} \prod_{i=1}^{j} \prod_{r=k-i+1 \vee 1}^{k-1} |(\sin \theta_r^{(i)}|^{r-1}$$

$$\iint_{\mathcal{U}\times\mathcal{K}} e^{-<e,s\perp(e+\sum_{i=1}^{j} t_i\, K H H_i(e_k)\,\Delta\, K H H_i(e_k)>} e^{2\sum_{i=1}^{j} \sqrt{t_i}\, <a\perp(m_i),\, s\perp K H H_i(e_k)>} s^{\rho\beta}\, d\mu(s)\, dK$$

(dK mesure de Haar sur $\mathcal{K}$).

Preuve : Noter qu'ici $x \rightsquigarrow s(x)$ a pour jacobien det T (en posant $s = \Pi(T\,{}^tT)$) soit encore $(\det({}^tTT))^{1/2}$, c'est-à-dire $(\det s^*)^{1/2}$ que nous notons $s^{*\frac{s(2\beta)}{2}} = s^{\beta}$.

Noter de plus que si dH est la mesure de Haar sur O(k) et si $(dH)_{\mathcal{K}}$ dénote (provisoirement) la mesure de Haar sur $\mathcal{K}$, on a l'égalité, pour toute fonction intégrable f :

$$\int_{O(k)} f(H)\, dH = \int_{O(k)/\mathcal{K}} d\bar{H} \int_{\mathcal{K}} f(KH)\, (dK)_{\mathcal{K}}$$

(la dernière intégrale ne dépend de H que par sa classe d'équivalence $\bar{H}$; pour la définition de $d\bar{H}$ voir [5] chapitre 15-18 par exemple).

THEOREME 10 :

Soit $f(t_1 \dots t_j)$ la valeur de l'intégrale écrite au théorème 9, sous l'hypothèse nulle. Le théorème 4 reste valable et généralise le test de Pillai.

Preuve : La loi des μ_i s'écrit sous la forme $dM(\rho_i)\, d\sigma(\mu_{1i})$ avec

$$\rho_i = ||\mu_i|| \quad , \quad \mu_{1i} = \frac{\mu_i}{\rho_i} \ .$$

La différentielle seconde étant calculée comme précédemment, on a :

$$E({}^t(a\perp(\mu_i')) \, A_{ii} \, a\perp(\mu_i) \mid \rho_1 \dots \rho_i) = \begin{cases} 0 \text{ si } i \neq i' \\ \frac{1}{k} \operatorname{tr} A_{ii} \, \rho_i^2 \quad \text{si } i = i' \end{cases}$$

Le test conduit à rejeter H_0 si :

$$\sum_{i=1}^{j} \langle e \, , A_{ii} \rangle > C_1 \qquad f$$

et le résultat s'obtient en explicitant A_{ii}.

VI - EXEMPLES ET CONTREXEMPLES

Pour être utilisé, ce théorème demande le calcul explicite de

$$f(t_1 \dots t_j) = \int_{\mathcal{K}} (e + \sum_{i=1}^{j} t_i \, K(V_i) \, \Delta \, K(V_i))^{\alpha - \rho\beta} \, dK$$

ce sont en fait des cas d'espèce.

Exemple 1 : Λ (donc Λ^{-1}) est diagonale. Il est facile de voir que la zone de rejet est :

$$\sum_{m=1}^{k} \frac{\sum_{i=1}^{j} X_{im}^2}{S_{mm} + \sum_{i=1}^{j} X_{im}^2} > C \ .$$

Exemple 2 : $\Lambda^{-1} \in \mathcal{M}_3^+$, et Λ^{-1} a un zéro en position (1,2). On a alors

$\alpha = (0,0,1) \quad \beta = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \quad \mathcal{K} = \{\text{identité}\}$, et

$$w^{\alpha - \rho\beta} = (w_{11} - \frac{w_{13}^2}{w_{33}})^{-\frac{\rho}{2}} (w_{22} - \frac{w_{33}^2}{w_{33}})^{-\frac{\rho}{2}} \; w_{33}^{\,1 - \frac{\rho}{2}} .$$

Le calcul de f est immédiat.

Exemple 3 : On suppose de plus que les deux premiers éléments diagonaux de Λ^{-1} sont égaux. On calcule $\alpha = (0,1)$, $\beta = (1, \frac{1}{2})$ et $\mathcal{K}$ est formé des rotations de type (1,2). De plus :

$$w^{\alpha-\rho\beta} = \left(\frac{w_{11} + w_{22}}{2} - \frac{w_{13}^2 + w_{23}^2}{2w_{33}}\right)^{-\rho} w_{33}^{1-\frac{\rho}{2}} .$$

Cette expression étant invariante par $\mathcal{K}$, le calcul de f est immédiat encore. Noter que même dans l'exemple 1, la loi de la statistique de rejet ne peut être obtenue explicitement.

Contrexemples : Nous avons supposé en tête du chapitre V que $\mathcal{P}$ était conservé par un sous groupe assez riche de transformations $u \rightsquigarrow {}^t g\ u\ g$. Par exemple, si $\mathcal{P}$ est de dimension $\frac{k(k+1)}{2}$ -1, donné par l'équation $\langle u,w\rangle = 0$, cela revient à supposer que w a 2 valeurs propres (exactement) non nulles (et de signes opposés !).

Pour k = 3 comment aborder l'étude de ce problème pour $\mathcal{P}$ donné, par exemple par $2\, a_{11}^* = a_{22}^* + a_{33}^*$?

$\mathcal{P} \cap \mathcal{M}_3^+$ est-il encore homogène ? (En autorisant des transformations linéaires de R^5 dans R^5 autres que celles ci-dessus), c'est peu probable. Si oui, quelle est l'interprétation physique de ces transformations ?

BIBLIOGRAPHIE

[1] BARRA J. R. - Notions fondamentales de statistique mathématique, Dunod - Paris (1971).

[2] FOURT G. - Lois β sur les cône homogènes, Annales scientifiques de l'Université de Clermont-Ferrand, N° 51 (1974) p. 22-30.

[3] FOURT G. - Mesures sur les cônes homogènes (non publié).

[4] FOURT G. - Au delà de l'analyse de la variance, Annales scientifiques de l'Université de Clermont-Ferrand, N° 65 (1977) p. 76-95.

[5] HEWITT E. et ROSS K.A. - Abstract harmonic analysis I, Springer Verlag - Berlin (1963).

[6] PILLAI K.C.S. et JAYACHANDRAN K. - Power comparisons of tests of two multivariate hypotheses based on four criteria, Biometrika 54 (1967) p. 195-210.

[7] SCHWARTZ R. - Local minimax test, Ann. Math. Statist. 38 (1967) p. 340-360.

[8] TAKEAKI KARIYA - The general MANOVA problem, Ann. of Statist. 6 (1978) p. 200-214.