G. LENNES

Inverses généralisés, problèmes d'estimation et épreuve de l'hypothèse linéaire générale en analyse statistique multivariate

Revue de statistique appliquée, tome 24, n° 3 (1976), p. 5-29

<http://www.numdam.org/item?id=RSA_1976__24_3_5_0>

INVERSES GÉNÉRALISÉS, PROBLÈMES D'ESTIMATION ET ÉPREUVE DE L'HYPOTHÈSE LINÉAIRE GÉNÉRALE EN ANALYSE STATISTIQUE MULTIVARIATE (1)

G. LENNES
Université de Liège – Belgique

1 – INTRODUCTION.

L'étude des modèles linéaires faisant intervenir des variables aléatoires scalaires peut être traitée comme un cas particulier de la régression linéaire. La méthode des moindres carrés conduit à un système d'équations linéaires en les paramètres : les équations normales. Dans le cas général de la régression, la matrice des coefficients de ce système est régulière ; par contre, elle est la plupart du temps singulière dans le cas particulier des modèles linéaires.

Afin d'éviter les difficultés que cela implique, les statisticiens ont été amenés à imposer des conditions supplémentaires aux paramètres de façon à obtenir un nouveau système d'équations normales qui soit régulier. Une telle "reparamétrisation" du modèle est toujours une procédure "ad hoc", variable selon le plan d'expérience considéré. En outre, comme le fait remarquer S.R. SEARLE, on peut trouver des conditions supplémentaires "pratiques" ou "évidentes" lorsqu'il s'agit d'un plan équilibré, mais cela n'est généralement pas possible dans le cas d'un plan non équilibré.

L'introduction des inverses généralisés a permis d'éviter ces difficultés et ainsi de fournir un traitement général des modèles linéaires applicable aussi bien aux plans équilibrés qu'aux autres. Les premiers résultats dans ce domaine sont dus aux travaux de R. PENROSE et C.R. RAO. Dans un ouvrage récent intitulé "Linear Models" S.R. SEARLE a publié un exposé systématique de la méthode et de ses applications, développant ainsi des thèmes déjà évoqués dans un livre précédent, "Matrix Algebra for the Biological Sciences".

Dans cet article, nous nous proposons d'étendre cette méthode au cas des variables aléatoires vectorielles. On retrouvera nécessairement dans notre travail des démarches voisines de celles adoptées par T.W. ANDERSON dans "An Introduction to Multivariate Statistical Analysis" et par S.R. SEARLE dans les deux ouvrages déjà cités. Il ne s'agit pas cependant d'une extension triviale de certains des résultats exposés par ces auteurs. Pour atteindre notre but, nous avons dû faire appel notamment à un théorème très général

---

(1) Article remis en Octobre 1975, révisé en Février 1976.

Mots-clés : **Inverses généralisés, Analyse statistique multivariate**

concernant les solutions d'équations matricielles en X de la forme AXB = C, théorème que l'on trouvera exposé dans "Generalized Inverse of Matrices and its Applications", par C.R. RAO et S.K. MITRA. En outre, nous effectuerons nos estimations à l'aide de la méthode du maximum de vraisemblance et non plus à l'aide de la méthode des moindres carrés. Nous devrons donc préciser dès le départ les distributions de nos variables aléatoires. Disons d'emblée que nous nous limiterons au cas des variables multidimensionnelles normales.

La suite de ce travail comporte quatre chapitres intitulés :

2 – Inverses généralisés d'une matrice

3 – Modèles linéaires – Présentation nouvelle

4 – Hypothèses linéaires – Présentation nouvelle

5 – Application.

2 – INVERSES GENERALISES D'UNE MATRICE.

Soit A une matrice rectangulaire à p lignes et q colonnes dont les éléments appartiennent au corps des nombres complexes ou des nombres réels. Nous désignerons par a_{ij} l'élément appartenant à la ligne de rang i et à la colonne de rang j et nous écrirons $A = (a_{ij})$.

On appelle inverse généralisé de A une matrice G à q lignes et p colonnes telle que :

$$AGA = A \tag{1}$$

Un tel inverse généralisé, ou g-inverse en abrégé, sera également désigné dans la suite par A^-. Si A est une matrice carrée régulière, son inverse au sens ordinaire, A^{-1}, satisfait à la relation (1).

Si outre la relation (1), la matrice G vérifie la relation

$$GAG = G \tag{2}$$

on dit que G constitue un g-inverse réflexif de A.

Nous allons maintenant rappeler quelques propriétés des g-inverses en nous limitant toutefois au corps des nombres réels.

T1 – Lorsque A est rectangulaire ou carrée singulière, il y a une infinité de matrices G satisfaisant à la relation (1). Elles s'obtiennent toutes à partir de l'une quelconque d'entre elles, A^-, grâce à la formule suivante :

$$G = A^-AA^- + U - A^-AUAA^-, \tag{3}$$

où U désigne une matrice quelconque à q lignes et p colonnes.

En faisant varier U, on engendrera l'ensemble des g-inverses de A.

Ce théorème n'est en fait qu'un cas particulier d'une proposition plus générale relative aux équations matricielles de la forme suivante.

T2 – Soit l'équation matricielle en X :

$$AXB = C \tag{4}$$

où A , X , B et C désignent respectivement des matrices p x q , q x r , r x s et p x s.

La condition nécessaire et suffisante pour que l'équation (4) admette une solution est que :

$$AA^{-}CB^{-}B = C \tag{5}$$

Cette condition satisfaite, la solution générale de (4) est

$$X = A^{-}CB^{-} + W - A^{-}AWBB^{-} \tag{6}$$

où W est une matrice q x r arbitraire.

T3 – Si nous désignons par A' la transposée de A, on a :

$$A\,(A'A)^{-}\,(A'A) = A \tag{7}$$

et

$$(A'A)\,(A'A)^{-}\,A' = A' \tag{7'}$$

T4 – Si A est symétrique, $(A^{-})'$ ainsi que $A^{-}A\,(A^{-})'$ et $(A^{-})'AA^{-}$ sont également des g-inverses de A.

Les démonstrations de ces diverses propositions se trouvent dans l'ouvrage de RAO et MITRA, "Generalized Inverse of Matrices and its Applications".

Dans le chapitre 4, nous considérerons des matrices K de format q x s vérifiant la relation suivante :

$$AA^{-}K = K, \tag{8}$$

où A désigne une matrice d'ordre q et A^{-} un g-inverse particulier de A. Voici quelques théorèmes relatifs à ces matrices K.

T5 – a) Rang K $\leqslant$ Rang A

b) $AA^{-}K = K$ entraîne $AGK = K$, où G désigne un g-inverse quelconque de A. En d'autres termes, la classe des matrices K vérifiant (8) ne dépend pas du choix d'un g-inverse particulier de A.

T6 – Si A est une matrice symétrique, les conditions $AA^{-}K = K$ et $K'A^{-}A = K'$ sont équivalentes.

T7 – Si A est une matrice symétrique et si $AA^{-}K = K, K'A^{-}K = K'GK$. En d'autres termes, la matrice $K'A^{-}K$ ne dépend pas du choix d'un g-inverse particulier de A.

De T4 et T7, on déduit aisément pour A symétrique.

T8 – a) $K'A^{-}A\,(A^{-})'K = K'\,(A^{-})'AA^{-}K = K'A^{-}K$ (9)

b) $K'A^{-}K$ est symétrique quel que soit le choix de A^{-}.

T9 – Si A est une matrice symétrique semi définie positive d'ordre q et de rang r et si K est une matrice de format q x s et de rang s, on a $s \leqslant r$ et la matrice $K'A^{-}K$ est régulière.

C'est là un théorème démontré par S.R. SEARLE dans "Matrix Algebra for the Biological Sciences".

Nous allons maintenant démontrer une propriété nouvelle.

T10 – Si A est une matrice symétrique semi définie positive d'ordre q et de rang r et si K est une matrice q x s de rang $s \leqslant r$ telle que $AA^-K = K$, il existe une matrice T non singulière d'ordre q telle que

$$A - K(K'A^-K)^{-1} K' = T\left(\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}\right)T' \qquad (10)$$

En particulier, si $s < r$, $A - K(K'A^-K)^{-1} K'$ est donc une matrice symétrique semi définie positive de rang $r - s$ et si $s = r$, cette matrice est nulle.

Démonstration

Compte tenu des hypothèses faites au sujet de A, il existe donc une matrice régulière F d'ordre q telle que

$$FAF' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad (11)$$

où I_r désigne la matrice unité d'ordre r.

On vérifie aisément que la matrice symétrique

$$A^- = F' \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} F \qquad (12)$$

constitue un g-inverse réflexif de A.

Puisque $(F')^{-1} = (F^{-1})'$ et compte tenu de (11), (12) et T9, on a :

$$A - K(K'A^-K)^{-1} K' = F^{-1} F(A - K(K'A^-K)^{-1} K') F' (F^{-1})'$$

$$= F^{-1}\left(\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} - FK\left(K'F'\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}FK\right)^{-1} K'F'\right)(F^{-1})'$$

$$= F^{-1}\left(\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} - S\left(S'\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}S\right)^{-1} S'\right)(F^{-1})' \qquad (13)$$

si on pose $S = FK$. (14)

S est donc une matrice q x s de rang s et compte tenu de (11), (12) et (14), la condition $AA^-K = K$ peut maintenant s'écrire :

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} S = S \qquad (15)$$

De cette dernière relation, on déduit que les $q - r$ dernières lignes de S sont nulles et que

$$S\left(S'\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}S\right)^{-1} S' = S(S'S)^{-1} S' \qquad (16)$$

Puisque les q – r dernières lignes sont nulles, posons

$$S = \begin{pmatrix} S_1 \\ 0 \end{pmatrix}, \text{ où } S_1 \text{ est une matrice r x s} \tag{17}$$

En remplaçant S par son expression détaillée dans le second membre de (16), il vient :

$$S\,(S'S)^{-1}\,S' = \begin{pmatrix} S_1\,(S_1'\,S_1)^{-1}\,S_1' & 0 \\ 0 & 0 \end{pmatrix} \tag{18}$$

La matrice $S_1\,(S_1'\,S_1)^{-1}\,S_1'$ est une matrice symétrique semi définie positive d'ordre r. On vérifie aisément qu'elle est idempotente.

Quant à sa trace, on a :

$$\operatorname{tr} S_1\,(S_1'\,S_1)^{-1}\,S_1' = \operatorname{tr} S_1'\,S_1\,(S_1'\,S_1)^{-1} = \operatorname{tr} I_s = s \tag{19}$$

Or, les valeurs propres d'une matrice idempotente sont toutes égales à 1 ou 0. En outre, la trace d'une telle matrice est égale au nombre de valeurs propres égales à 1.

Compte tenu de (19), la matrice $S_1\,(S_1'\,S_1)^{-1}\,S_1'$ aura s valeurs propres égales à 1 et r – s valeurs propres égales à 0.

Il existe donc une matrice othogonale U_1 d'ordre r telle que :

$$U_1\,S_1\,(S_1'\,S_1)^{-1}\,S_1'\,U_1' = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \tag{20}$$

Considérons maintenant la matrice U d'ordre q définie comme suit :

$$U = \begin{pmatrix} U_1 & 0 \\ 0 & I_{q-r} \end{pmatrix} \tag{21}$$

On vérifie aisément que U est orthogonale et que

$$U\,S\,(S'\,S)^{-1}\,S'\,U' = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \tag{22}$$

$$U \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} U' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \tag{23}$$

Compte tenu des formules (16), (22), (23) et du fait que U est orthogonale, la relation (13) peut s'écrire :

$$\begin{aligned} A - K\,(K'A^-K)^{-1}\,K' &= F^{-1}\,U' \left(\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \right) U\,(F^{-1})' \\ &= T \left(\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \right) T' \end{aligned} \tag{24}$$

si on pose $T = F^{-1}\,U'$

3 – MODELES LINEAIRES – PRESENTATION NOUVELLE.

1 – Généralités.

Reprenant les notations de T.W. ANDERSON, nous représenterons par $\vec{x}_1, \ldots, \vec{x}_N$ un ensemble de N vecteurs observés, $\vec{x}_\alpha$ ayant été extrait au hasard d'une distribution normale de moyenne $B\vec{z}_\alpha$ et de matrice des covariances Σ , α variant de 1 à N.

Nous désignerons plus brièvement de telles distributions par $N(B\vec{z}_\alpha, \Sigma)$. Les densités de probabilité qui les définissent peuvent s'écrire comme suit :

$$| \Sigma^{-1} |^{\frac{1}{2}} (2\pi)^{-\frac{p}{2}} e^{-\frac{1}{2}(\vec{x}_\alpha - B\vec{z}_\alpha)' \Sigma^{-1} (\vec{x}_\alpha - B\vec{z}_\alpha)} \tag{25}$$

Σ (et par suite Σ^{-1}) est une matrice symétrique définie positive d'ordre p,B est une matrice de format p x q, tandis que $\vec{x}_\alpha$ et $\vec{z}_\alpha$ sont des vecteurs colonnes ayant respectivement p et q composantes :

$$\vec{x}_\alpha = \begin{pmatrix} x_{1\alpha} \\ \vdots \\ x_{p\alpha} \end{pmatrix} \text{ et } \vec{z}_\alpha = \begin{pmatrix} z_{1\alpha} \\ \vdots \\ z_{q\alpha} \end{pmatrix}, \; (\alpha = 1, \ldots, N) \tag{26}$$

Nous représenterons par X la matrice p x N dont les N colonnes sont les vecteurs $\vec{x}_\alpha$ et par Z la matrice q x N dont les N colonnes sont les vecteurs $\vec{z}_\alpha$. Ces $\vec{z}_\alpha$ sont des vecteurs constants donnés, ils définissent la structure du plan d'expérience. T.W. ANDERSON a limité son étude aux cas où le rang de Z est égal à q. En outre, le nombre N d'observations est choisi de façon que $N \geqslant p + q$.

Nous nous placerons ici dans le cas général où le rang de Z est égal à r et nous traiterons plus particulièrement le cas où $r < q$.

Nous désignerons respectivement par b_{ik}, σ_{ij} et σ^{ij} les éléments des matrices B , Σ et Σ^{-1} :

$$B = (b_{ik}) , \Sigma = (\sigma_{ij}) , \Sigma^{-1} = (\sigma^{ij}) \tag{27}$$

Ces éléments sont les paramètres du modèle.

Afin de ne pas multiplier les notations, nous représenterons par $\vec{x}_\alpha$ non seulement une observation mais aussi la variable aléatoire correspondant à $N(B\vec{z}_\alpha, \Sigma)$. Le contexte indiquera toujours clairement le sens qu'il faut attribuer aux symboles. Ainsi, lorsque nous parlerons de l'espérance mathématique de $\vec{x}_\alpha$, en abrégé $\mathcal{E}\, \vec{x}_\alpha$, c'est évidemment de la variable aléatoire $\vec{x}_\alpha$ qu'il s'agira.

Avec cette convention, le modèle linéaire destiné à rendre compte des observations s'écrira :

$$\mathcal{E}\,(X) = B\,Z \tag{28}$$

On notera encore que les variables aléatoires normales $\vec{x}_\alpha$ et $\vec{x}_\beta$ sont indépendantes lorsque $\alpha \neq \beta$:

$$\mathrm{Cov}\,(\vec{x}_\alpha, \vec{x}_\beta) = \mathcal{E}\,(\vec{x}_\alpha - B\vec{z}_\alpha)\,(\vec{x}_\beta - B\vec{z}_\beta)' = \delta_{\alpha\beta}\,\Sigma \tag{29}$$

Les matrices des paramètres B et Σ sont des matrices fixes dont les éléments nous sont inconnus.

Pour estimer les matrices B et Σ, on utilisera la méthode du maximum de vraisemblance. Compte tenu de (25), le logarithme népérien de la vraisemblance L de l'échantillon est donné par l'expression suivante :

$$\log_e L = -\frac{1}{2} Np \log_e 2\pi + \frac{1}{2} N \log_e |\Sigma^{-1}| - \frac{1}{2} \mathrm{tr}\,(\Sigma^{-1} D), \tag{30}$$

où
$$D = \sum_{\alpha=1}^{N} (\vec{x}_\alpha - B\vec{z}_\alpha)\,(\vec{x}_\alpha - B\vec{z}_\alpha)' = (X - BZ)\,(X - BZ)' \tag{31}$$

Dans ces expressions, B et Σ ne désignent plus les matrices des paramètres de notre modèle linéaire, mais bien des matrices arbitraires telles que $\log_e L$ ait un sens : Σ (et par suite Σ^{-1}) sera une matrice définie positive quelconque.

Les équations du maximum de vraisemblance s'obtiennent en annulant les dérivées partielles de $\log_e L$ par rapport aux éléments b_{rs} et σ^{rs} des matrices B et Σ^{-1}.

Posons
$$A = \sum_{\alpha=1}^{N} \vec{z}_\alpha\, \vec{z}_\alpha{}' = ZZ' \tag{32}$$

$$C = \sum_{\alpha=1}^{N} \vec{x}_\alpha\, \vec{z}_\alpha{}' = XZ' \tag{33}$$

Sous forme matricielle, le système des équations du maximum de vraisemblance s'écrira alors comme suit :

$$\begin{cases} B\,A = C & (34) \\ N\,\Sigma = D = (X - BZ)\,(X - BZ)' & (35) \end{cases}$$

2 – Résolution du système des équations du maximum de vraisemblance.

a) *Cas de B*

Il s'agit de résoudre la première des équations du système. Mais ici, rang A = rang ZZ' = rang Z = $r < q$. La matrice A d'ordre q n'est plus régulière, contrairement à ce qui se passait dans le cas étudié par T.W. ANDERSON. Pour traiter notre problème, nous aurons recours aux inverses généralisés.

L'équation BA = C est du type AXB = C. Nous devrons tout d'abord nous demander si elle admet des solutions ou, en d'autres termes, si la condition (5) de T2 est satisfaite.

Avec les notations actuelles, cette condition s'écrit

$$CA^{-}A = C, \tag{36}$$

ou plus explicitement :

$$XZ' (ZZ')^{-}(ZZ') = XZ' \tag{37}$$

Or, l'identité (7) du chapitre 2 peut s'écrire :

$$Z' (ZZ')^{-}(ZZ') = Z' \tag{38}$$

La relation (37) n'est rien d'autre que l'identité (38) dont les deux membres ont été multipliés à gauche par X. Elle sera donc toujours vérifiée et l'équation BA = C aura toujours des solutions. Une solution quelconque $\hat{\hat{B}}$ sera donnée par la formule suivante :

$$\hat{\hat{B}} = CA^{-} + W - WAA^{-} = CA^{-} + W (I - AA^{-}), \tag{39}$$

où W désigne une matrice p x q arbitraire.

Si nous prenons W = 0, nous obtenons la solution particulière :

$$\hat{B} = CA^{-}, \tag{40}$$

d'où

$$\hat{\hat{B}} = \hat{B} + W (I - AA^{-}) \tag{41}$$

b) *Cas de* Σ

Nous allons voir que si les équations du maximum de vraisemblance ne déterminent pas B univoquement, elles vont nous conduire par contre à une solution unique pour Σ.

Dans la seconde des équations du système, remplaçons respectivement D par $\hat{D} = (X - \hat{B}Z)\ (X - \hat{B}Z)'$ et par $\hat{\hat{D}} = (X - \hat{\hat{B}}Z)\ (X - \hat{\hat{B}}Z)'$. On obtient alors :

$$\hat{\Sigma} = N^{-1}\,\hat{D} \tag{42}$$

et

$$\hat{\hat{\Sigma}} = N^{-1}\,\hat{\hat{D}} \tag{42'}$$

Compte tenu de (41), on a :

$$X - \hat{\hat{B}}Z = X - \hat{B}Z - W (I - AA^{-})\, Z \tag{43}$$

Or, T3 du chapitre 2 montre que

$$(I - AA^{-})\, Z = (I - (ZZ')\,(ZZ')^{-})\, Z = 0 \tag{44}$$

On en déduit donc que

$$X - \hat{\hat{B}}Z = X - \hat{B}Z \tag{45}$$

et par suite :

$$\hat{\hat{D}} = \hat{D} \tag{46}$$

$$\hat{\hat{\Sigma}} = \hat{\Sigma} \tag{46'}$$

On démontrera dans le § 5 de ce chapitre que la matrice symétrique $\hat{D}$ (et par suite $\hat{\Sigma}$) est définie positive (voir T13).

3 – Le maximum de $\log_e L$

T11 – Quelle que soit la solution choixie du système des équations du maximum de vraisemblance, $(\hat{B}, \hat{\Sigma})$ ou $(\hat{\hat{B}}, \hat{\hat{\Sigma}})$, la valeur prise par $\log_e L$ est la même.

En effet, la formule (30) permettant de calculer $\log_e L$ ne fait intervenir explicitement, outre des constantes, que Σ et D. Or, nous venons précisément de voir que $\hat{\hat{\Sigma}} = \hat{\Sigma}$ et $\hat{\hat{D}} = \hat{D}$.

Les équations du maximum de vraisemblance constituent simplement des conditions nécessaires à l'existence d'un extremum. Il s'agira maintenant de démontrer le théorème suivant :

T12 – $(\log_e L)_{\hat{B},\hat{\Sigma}}$ constitue le maximum de $\log_e L$.

Démonstration

a) Démontrons tout d'abord que quelle que soit Σ définie positive et quelle que soit B telle que D soit définie positive, on a :

$$(\log_e L)_{\hat{B},\Sigma} - (\log_e L)_{B,\Sigma} \geqslant 0 \tag{47}$$

De l'expression (30) de $\log_e L$, on tire :

$$(\log_e L)_{\hat{B},\Sigma} - (\log_e L)_{B,\Sigma} = \frac{1}{2} \operatorname{tr} (\Sigma^{-1} (D - \hat{D})) \tag{48}$$

On peut toujours écrire :

$$X - BZ = X - \hat{B}Z + (\hat{B} - B) Z \tag{49}$$

Compte tenu du fait que $\hat{B}A = C$, on en déduit aisément que

$$D - \hat{D} = (\hat{B} - B) ZZ' (\hat{B} - B)' \tag{50}$$

Cette dernière formule montre que $D - \hat{D}$ est semi définie positive, il en va de même de Σ^{-1} par hypothèse. Il suit donc que :

$$\operatorname{tr} (\Sigma^{-1} (D - \hat{D})) \geqslant 0 \tag{51}$$

De (48) et (51), on déduit enfin la formule (47).

b) Démontrons de même que quelles que soient Σ et B satisfaisant aux conditions précédentes, on a :

$$(\log_e L)_{B,N^{-1}D} - (\log_e L)_{B,\Sigma} \geqslant 0 \tag{52}$$

De l'expression (30) de $\log_e L$, on déduit :

$$(\log_e L)_{B,N^{-1}D} - (\log_e L)_{B,\Sigma}$$

$$= \frac{N}{2} \log_e | ND^{-1} | - \frac{1}{2} \mathrm{tr}\,(NI_p) - \frac{N}{2} \log_e | \Sigma^{-1} | + \frac{1}{2} \mathrm{tr}\,(\Sigma^{-1} D)$$

$$= \frac{N}{2} (- \log_e | N^{-1} \Sigma^{-1} D | - p + \mathrm{tr}\,(N^{-1} \Sigma^{-1} D)) \tag{53}$$

$N^{-1} \Sigma^{-1}$ et D étant définies positives, il existe donc une matrice régulière T telle que :

$$T' N^{-1} \Sigma^{-1} T = \mathrm{diag}\,(s_1, \ldots, s_p), \text{ avec } s_i > 0 \tag{54}$$

$$(T^{-1}) D (T^{-1})' = \mathrm{diag}\,(d_1, \ldots, d_p), \text{ avec } d_i > 0 \tag{55}$$

De (54) et (55), on déduit aisément que :

$$T' N^{-1} \Sigma^{-1} D (T')^{-1} = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_p), \tag{56}$$

avec
$$\lambda_i = s_i d_i > 0, (i = 1, \ldots, p) \tag{57}$$

Or,

$$\log_e | N^{-1} \Sigma^{-1} D | = \log_e | T' N^{-1} D (T')^{-1} | = \sum_{i=1}^{p} \log_e \lambda_i \tag{58}$$

et
$$\mathrm{tr}\,(N^{-1} \Sigma^{-1} D) = \mathrm{tr}\,(T' N^{-1} \Sigma^{-1} D (T')^{-1}) = \sum_{i=1}^{p} \lambda_i \tag{59}$$

En tenant compte de (58) et (59), la relation (53) peut s'écrire :

$$(\log_e L)_{B,N^{-1}D} - (\log_e L)_{B,\Sigma} = \frac{N}{2} \sum_{i=1}^{p} (- \log_e \lambda_i - 1 + \lambda_i) \tag{60}$$

Mais, pour tout λ positif, on a :

$$\lambda - 1 - \log_e \lambda \geqslant 0 \tag{61}$$

L'inégalité (61) est applicable à chacun des p termes du second membre de (60), puisque $\lambda_i > 0$ $(i = 1, \ldots, p)$. On en déduit l'inégalité (52).

c) Dans l'inégalité (52), remplaçons B par $\hat{B}$; nous devrons alors remplacer $N^{-1} D$ par $N^{-1} \hat{D} = \hat{\Sigma}$. Nous avons alors :

$$(\log_e L)_{\hat{B},\hat{\Sigma}} \geqslant (\log_e L)_{\hat{B},\Sigma} \tag{62}$$

D'autre part, de l'inégalité (47), on déduit :

$$(\log_e L)_{\hat{B},\Sigma} \geqslant (\log_e L)_{B,\Sigma} \tag{63}$$

De (62) et (63), il suit :

$$(\log_e L)_{\hat{B},\hat{\Sigma}} \geqslant (\log_e L)_{B,\Sigma} \tag{64}$$

4 – Distribution d'échantillonnage de $\hat{B}$

Des formules (28) et (40), on déduit :

$$\mathcal{E}(\hat{B}) = \mathcal{E}(CA^-) = \mathcal{E}(XZ'(ZZ')^-) = \mathcal{E}(X)\,Z'(ZZ')^-$$
$$= BZZ'(ZZ')^- = BAA^- \tag{65}$$

De même, de (28) et (41), il suit :

$$\mathcal{E}(\hat{\hat{B}}) = \mathcal{E}(\hat{B}) + W(I - AA^-) \tag{66}$$

Ainsi, ni $\hat{B}$, ni $\hat{\hat{B}}$ ne constituent des estimations sans biais de B. Nous verrons cependant plus loin que certaines fonctions BK de B sont estimables en un sens que nous préciserons alors.

On constate aisément que :

$$\hat{B} - \mathcal{E}(\hat{B}) = \hat{\hat{B}} - \mathcal{E}(\hat{\hat{B}}) \tag{67}$$

Dans la suite, nous nous occuperons exclusivement de $\hat{B} - \mathcal{E}(\hat{B})$.

$$\hat{B} - \mathcal{E}(\hat{B}) = (C - \mathcal{E}(C))\,A^- = (X - \mathcal{E}(X))\,Z'\,A^- \tag{68}$$

Désignons par $\vec{\hat{b}}_i$ et $\vec{x}_i$ les lignes de rang i des matrices $\hat{B}$ et X. Il s'agit là de vecteurs lignes ayant respectivement q et N composantes

Nous allons rechercher la matrice des covariances de $\vec{\hat{b}}_i$ et $\vec{\hat{b}}_j$.

Compte tenu de (68), on aura donc :

$$\vec{\hat{b}}_i - \mathcal{E}(\vec{\hat{b}}_i) = (\vec{x}_i - \mathcal{E}(\vec{x}_i))\,Z'\,A^- \tag{69}$$

D'où,
$$\mathrm{Cov}(\vec{\hat{b}}_i, \vec{\hat{b}}_j) = \mathcal{E}\,\{(\vec{\hat{b}}_i - \mathcal{E}(\vec{\hat{b}}_i))'\,(\vec{\hat{b}}_j - \mathcal{E}(\vec{\hat{b}}_j))\}$$
$$= (A^-)'\,Z\,\mathcal{E}\,\{(\vec{x}_i - \mathcal{E}(\vec{x}_i))'\,(\vec{x}_j - \mathcal{E}(\vec{x}_j))\}\,Z'\,A^- \tag{70}$$

De la formule (29) selon laquelle $\mathrm{Cov}(\vec{x}_\alpha, \vec{x}_\beta) = \delta_{\alpha\beta}\,\Sigma$, on déduit aisément que :

$$\mathcal{E}\{(\vec{x}_i - \mathcal{E}(\vec{x}_i))'\,(\vec{x}_j - \mathcal{E}(\vec{x}_j))\} = \sigma_{ij}\,I_N \tag{71}$$

(On notera que $\vec{x}_i$ avec un indice latin désigne la ligne de rang i de X et $\vec{x}_\alpha$ avec un indice grec, la colonne de rang α).

Remplaçant dans (70), il vient :

$$\mathrm{Cov}(\vec{\hat{b}}_i, \vec{\hat{b}}_j) = (A^-)'\,Z\,\sigma_{ij}\,I_N\,Z'\,A^- = \sigma_{ij}\,(A^-)'\,AA^- \tag{72}$$

5 – Distribution d'échantillonnage de $\hat{\Sigma}$.

Si on développe les expressions définissant $\hat{D}$ et $\hat{\hat{D}}$ en tenant compte de ce que $\hat{B}A = \hat{\hat{B}}A = C$, on obtient aisément les relations suivantes :

$$N\,\hat{\Sigma} = \hat{D} = XX' - \hat{B}A\hat{B}' \tag{73}$$

$$N\,\hat{\hat{\Sigma}} = \hat{\hat{D}} = XX' - \hat{\hat{B}}A\hat{\hat{B}}' \tag{74}$$

Comme $\hat{D} = \hat{\hat{D}}$, on en déduit :

$$\hat{B}A\hat{B}' = \hat{\hat{B}}A\hat{\hat{B}}' \tag{75}$$

Ces formules sont à la base de l'étude de la distribution d'échantillonnage de $N\,\hat{\Sigma}$. Elles vont nous permettre de démontrer la proposition suivante qui étend le théorème 8.2.2. de T.W. ANDERSON (log. cit., p. 183) au cas où $r = \text{Rang } Z < q$:

T13 – a) $\hat{\Sigma}$ est distribuée indépendamment de $\hat{B}$

b) $N\,\hat{\Sigma}$ obéit à une distribution de WISHART $W(\Sigma\,, N - r)$, où N désigne l'effectif de l'échantillon et r le rang de Z.

Ainsi, $\dfrac{N\,\hat{\Sigma}}{N-r} = \dfrac{\hat{D}}{N-r}$ constituera une estimation sans biais de Σ.

Démonstration

1) Puisque $A = ZZ'$ est une matrice symétrique semi définie positive de rang r, il existe une matrice régulière F d'ordre q telle que :

$$FZZ'\,F' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = E_0\,E_0', \tag{76}$$

si nous posons $E_0 = FZ$ (77)

E_0 est une matrice à q lignes et N colonnes. La relation (76) montre que les r premières lignes sont des vecteurs orthogonaux deux à deux et de longueur égale à l'unité, tandis que les $q - r$ dernières lignes sont formées uniquement de zéros :

$$E_0 = \begin{pmatrix} E_1 \\ 0 \end{pmatrix}, \tag{78}$$

avec $E_1\,E_1' = I_r$ (79)

Les r vecteurs lignes orthonormés formant E_1 appartiennent à l'espace vectoriel de dimension N. Il est bien connu qu'on peut leur adjoindre $N - r$ autres vecteurs lignes, formant une nouvelle matrice que nous désignerons par E_2, de sorte que l'ensemble des N vecteurs constitue une base orthonormée de l'espace vectoriel à N dimensions.

En d'autres termes, si nous posons :

$$E = (e_{\alpha\beta}) = \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}, \tag{80}$$

la matrice E d'ordre N ainsi définie est orthogonale :

$$EE' = E'\,E = I_N \tag{81}$$

Définissons maintenant N vecteurs $\vec{u}_\alpha$:

$$\vec{u}_\alpha = \sum_{\beta=1}^{N} e_{\alpha\beta}\, \vec{x}_\beta\ , \ (\alpha = 1, \dots, N) \tag{82}$$

Puisque E est orthogonale et que $\vec{x}_1, \dots, \vec{x}_N$ sont N variables normales indépendantes de même matrice des covariances Σ, il en va de même de $\vec{u}_1, \dots, \vec{u}_N$.

Désignons par U la matrice p x N dont les N colonnes sont les vecteurs $\vec{u}_\alpha$. Les relations (82) peuvent s'écrire :

$$U = XE' \tag{83}$$

A partir de (81) et (83), on démontre aisément que :

$$\sum_{\alpha=1}^{N} \vec{u}_\alpha\, \vec{u}_\alpha' = UU' = XE'\, EX' = XX' = \sum_{\alpha=1}^{N} \vec{x}_\alpha\, \vec{x}_\alpha' \tag{84}$$

Enfin des formules (73) et (84), on déduit :

$$\sum_{\alpha=1}^{N} \vec{u}_\alpha\, \vec{u}_\alpha' = N\,\hat{\Sigma} + \hat{B}A\hat{B}' \tag{85}$$

2) Nous allons maintenant rechercher l'expression de $\hat{B}A\hat{B}'$ en fonction des $\vec{u}_\alpha$. Considérons $\hat{B} = CA^-$, où A^- est le g-inverse réflexif et symétrique défini comme suit :

$$A^- = F' \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} F, \tag{86}$$

F vérifiant la relation (76). On a donc :

$$\hat{B}A\hat{B}' = CA^-\, C' = XZ'\, F' \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} FZX' \tag{87}$$

Des formules (81) et (83), on déduit :

$$X = UE \tag{88}$$

En remplaçant X par UE et FZ par E_0 dans le second membre de (87) et en utilisant les relations (78) (80) et (81), on obtient :

$$\hat{B}A\hat{B}' = UE_0' \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} E_0\, E'\, U' = \sum_{\alpha=1}^{r} \vec{u}_\alpha\, \vec{u}_\alpha' \tag{89}$$

Enfin, de (85) et (89), il vient :

$$N\,\hat{\Sigma} = \sum_{\alpha=r+1}^{N} \vec{u}_\alpha\, \vec{u}_\alpha' \tag{90}$$

Les formules (89) et (90) démontrent le (a) de T13 puisque $\vec{u}_\alpha$ est indépendant de $\vec{u}_\beta$ lorsque $\alpha \neq \beta$.

3) Pour obtenir le (b) de T13, il suffit de démontrer que $\mathcal{E}(\vec{u}_\alpha) = 0$ pour $\alpha = r + 1, \ldots, N$.

Des formules (28) et (83), on déduit :

$$\mathcal{E}(U) = \mathcal{E}(X)\, E' = BZE' \tag{91}$$

De (77), on tire, puisque F est régulière

$$Z = F^{-1} E_0 \tag{92}$$

En remplaçant dans (91) et en tenant compte des relations (78), (80) et (81), il vient :

$$\mathcal{E}(U) = BF^{-1} E_0 E' = BF^{-1} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \tag{93}$$

4 – HYPOTHESES LINEAIRES – PRESENTATION NOUVELLE.

1 – Fonctions invariantes – Fonctions estimables.

Désignons par K une matrice constante de format q x s.

On peut se demander s'il n'existe pas des fonctions matricielles de la forme BK et qui soient telles que : $\hat{\hat{B}}K = \hat{B}K$ (94)

En d'autres termes, on recherche des fonctions de B qui prendraient la même valeur pour les diverses solutions de l'équation BA = C.

Compte tenu de l'expression (41) de $\hat{\hat{B}}$, on peut formuler cette exigence comme suit :

$$\hat{B}K + W(I - AA^-)K = \hat{B}K \tag{95}$$

où,

$$W(I - AA^-)K = 0, \tag{96}$$

quelle que soit la matrice W de format p x q.

On voit donc qu'il suffit de choisir K de sorte que $AA^- K = K$. C'est là la condition (8), qui a déjà été étudiée dans le chapitre 2. On peut également montrer que cette condition suffisante est aussi nécessaire.

Considérons maintenant une fonction BK telle que $AA^- K = K$. Compte tenu des formules (8), (28) et (41), on a :

$$\mathcal{E}(\hat{\hat{B}}K) = \mathcal{E}(\hat{B}K) = \mathcal{E}(XZ' A^- K) = BAA^- K = BK \tag{97}$$

Ainsi, lorsque BK est une fonction matricielle invariante, BK en constitue une estimation sans biais. Les fonctions matricielles invariantes sont donc estimables.

La ligne de rang i de $\hat{B}K$ s'obtient en calculant $\vec{\hat{b}}_j K$ où $\vec{\hat{b}}_i$ désigne la ligne de rang i de B. Compte tenu des formules (9) et (72), on peut démontrer que :

$$\begin{aligned} \mathrm{Cov}\,(\vec{\hat{b}}_i K\,,\vec{\hat{b}}_j K) &= \mathcal{E}\,\{K'\,(\vec{\hat{b}}_i - \mathcal{E}\,(\vec{\hat{b}}_i))'\,(\vec{\hat{b}}_j - \mathcal{E}\,(\vec{\hat{b}}_j))\,K\} \\ &= \sigma_{ij}\,K'\,(A^-)'\,AA^-\,K = \sigma_{ij}\,K'\,A^-\,K \end{aligned} \tag{98}$$

2 – Hypothèse linéaire et modèle réduit.

Aux conditions que $\vec{x}_\alpha$ provienne de N $(B\vec{z}_\alpha\,,\Sigma)$ et que rang Z = r, adjoignons maintenant l'hypothèse linéaire :

$$(H) : BK = M, \tag{99}$$

où M est une matrice constante donnée et K une matrice q x s satisfaisant à la condition $AA^-K = K$ et que nous supposerons en outre de rang s. Rappelons que l'on a nécessairement $s \leqslant r$, comme le montre T5 du chapitre 2.

Nous obtenons ainsi un nouveau modèle que l'on qualifie de réduit.

On va rechercher des matrices B et Σ rendant $\log_e L$ maximum, compte tenu de la condition BK = M, K étant une matrice vérifiant (8). Il s'agit là d'un problème d'extremum lié que nous avons traité par la méthode des multiplicateurs de Lagrange.

Posons
$$V = (v_{ih}) = BK - M \tag{100}$$

Nous sommes conduit à rechercher le maximum de

$$\varphi = \log_e L + \sum_{i=1}^{p} \sum_{h=1}^{s} v_{ih}\,t_{ih}, \tag{101}$$

où $(t_{ih}) = T$ est une matrice p x s dont les éléments t_{ih} sont les multiplicateurs de Lagrange.

Les équations obtenues en annulant les dérivées partielles de φ par rapport aux éléments $b_{ru}\,,t_{ru}$ et σ^{ru} des matrices B, T et Σ^{-1} peuvent se résumer sous forme matricielle de la façon suivante :

$$\left\{\begin{array}{ll} C - BA + \Sigma\,TK' = 0 & (102) \\ BK = M & (103) \\ N\,\Sigma = D = (X - BZ)\,(X - BZ)' & (104) \end{array}\right.$$

3 – Résolution du système des équations du maximum de vraisemblance.

a) *Cas de B*

Désignons par $\approx\atop{B}$ une solution quelconque de la première équation du système. On peut toujours écrire :

$$\overset{\approx}{B} = \hat{B} + W, \tag{105}$$

où W sera une matrice convenablement choisie.

Remplaçons B par $\approx{B}$ dans la première équation du système. Compte tenu de ce que $C - \hat{B}A = 0$, on voit que W doit satisfaire l'équation matricielle :

$$WA = \Sigma\ TK' \tag{106}$$

La condition nécessaire et suffisante pour que l'équation (106) admette des solutions, soit la relation (5) du chapitre 2, s'écrit ici :

$$\Sigma\ TK'\ A^{-}A = \Sigma\ TK' \tag{107}$$

Elle sera toujours satisfaite. En effet, K vérifie par hypothèse la relation $AA^{-}K = K$. Or, puisque A est symétrique, T6 nous permet d'affirmer qu'elle est équivalente à $K'\ A^{-}A = K'$.

D'après T2, une solution quelconque de (106) peut s'exprimer comme suit :

$$W = \Sigma\ TK'\ A^{-} + U\ (I - AA^{-}), \tag{108}$$

où U est une matrice arbitraire p x q. Ainsi,

$$\overset{\approx}{B} = \hat{B} + \Sigma\ TK'\ A^{-} + U\ (I - AA^{-}) \tag{109}$$

Remplaçons B par $\overset{\approx}{B}$ dans la seconde équation du système. Compte tenu de (8), on obtient :

$$\Sigma\ TK'\ A^{-}K = -\ (\hat{B}K - M) \tag{110}$$

Puisque K est de rang s, T9 permet d'en déduire :

$$\Sigma\ T = -\ (\hat{B}K - M)\ (K'\ A^{-}K)^{-1} \tag{111}$$

Une solution quelconque du système formé par les deux premières équations du système s'exprimera donc comme suit :

$$\overset{\approx}{B} = \hat{B} - (\hat{B}K - M)\ (K'\ A^{-}K)^{-1}\ K'\ A^{-} + U\ (I - AA^{-}) \tag{112}$$

Ainsi, dans le cas du modèle réduit également, la méthode du maximum de vraisemblance ne conduit pas à une solution unique : $\overset{\approx}{B}$ dépend d'une matrice arbitraire U.

Désignons par $\tilde{B}$ la solution particulière correspondant à $U = 0$:

$$\tilde{B} = \hat{B} - (\hat{B}K - M)\ (K'\ A^{-}K)^{-1}\ K'\ A^{-} \tag{113}$$

On a :

$$\overset{\approx}{B} = \tilde{B} + U\ (I - AA^{-}) \tag{114}$$

b) *Cas de* Σ

Nous allons voir également que si les équations du maximum de vraisemblance ne déterminent pas B univoquement, elles vont nous conduire par contre à une solution unique pour Σ.

Dans la troisième des équations du système, remplaçons D respectivement par $\widetilde{D} = (X - \widetilde{B}Z)(X - \widetilde{B}Z)'$ et par $\widetilde{\widetilde{D}} = (X - \widetilde{\widetilde{B}}Z)(X - \widetilde{\widetilde{B}}Z)'$.

On obtient alors :

$$\widetilde{\Sigma} = N^{-1}\widetilde{D} \tag{115}$$

et

$$\widetilde{\widetilde{\Sigma}} = N^{-1}\widetilde{\widetilde{D}} \tag{115'}$$

Grâce à T3, on déduit aisément de (114) que :

$$X - \widetilde{\widetilde{B}}Z = X - \widetilde{B}Z \tag{116}$$

et par suite :

$$\widetilde{\widetilde{D}} = \widetilde{D} \tag{117}$$

$$\widetilde{\widetilde{\Sigma}} = \widetilde{\Sigma} \tag{117'}$$

Nous allons maintenant rechercher une expression commode de $\widetilde{D}$.

De la formule (113), on tire :

$$X - \widetilde{B}Z = X - \hat{B}Z + (\hat{B}K - M)(K'A^{-}K)^{-1}K'A^{-}Z \tag{118}$$

Enfin, on démontre aisément que :

$$\widetilde{D} = \hat{D} + (\hat{B}K - M)(K'A^{-}K)^{-1}(\hat{B}K - M)', \tag{119}$$

si dans le développement de $(X - \widetilde{B}Z)(X - \widetilde{B}Z)'$, on tient compte de ce que $\hat{B}A = C$ et de T8 du chapitre 2.

Remarquons enfin que $\widetilde{D}$ (et par suite $\widetilde{\Sigma}$) est une matrice symétrique définie positive, comme le montre (119).

4 – Le maximum lié de $\log_e L$.

T14 – Quelle que soit la solution choisie du système des équations du maximum de vraisemblance, la valeur prise par $\log_e L$ est le même.

Ce théorème est une conséquence immédiate de (30), (117) et (117').

T15 – Quelles que soient Σ définie positive et B satisfaisant à $BK = M$, on a :

$$(\log_e L)_{\widetilde{B},\widetilde{\Sigma}} \geqslant (\log_e L)_{B,\Sigma} \tag{120}$$

Démonstration

a) Démontrons tout d'abord que quelles que soient B et Σ satisfaisant aux conditions du théorème, on a :

$$(\log_e L)_{\widetilde{B},\Sigma} - (\log_e L)_{B,\Sigma} \geqslant 0 \tag{121}$$

De l'expression (30) de $\log_e L$, on tire :

$$(\log_e L)_{\widetilde{B},\Sigma} - (\log_e L)_{B,\Sigma} = \frac{1}{2}\operatorname{tr}(\Sigma^{-1}(D - \widetilde{D})) \tag{122}$$

De la formule (119), on déduit :

$$D - \widetilde{D} = (D - \hat{D}) - (\hat{B}K - M)(K' A^{-}K)^{-1} (\hat{B}K - M)' \tag{123}$$

Dans le second membre de (123), remplaçons $D - \hat{D}$ par l'expression donnée dans la formule (50) du chapitre 3 et remplaçons M par BK. On obtient :

$$D - \widetilde{D} = (\hat{B} - B)(A - K(K' A^{-}K)^{-1} K')(\hat{B} - B)' \tag{124}$$

Or, d'après T10 du chapitre 2, $A - K(K' A^{-}K)^{-1} K'$, est une matrice symétrique semi définie positive. La formule (124) montre qu'il en va de même de $D - \widetilde{D}$. Dès lors,

$$\mathrm{tr}\,(\Sigma^{-1}(D - \widetilde{D})) \geqslant 0 \tag{125}$$

De (122) et (125), on déduit (121).

b) Dans le point (b) de la démonstration de T12, on a démontré dans le chapitre 3 l'inégalité (52), soit :

$$(\log_e L)_{B,N^{-1}D} - (\log_e L)_{B,\Sigma} \geqslant 0$$

Elle est évidemment encore valable si on suppose BK = M.

c) Dans l'inégalité (52), remplaçons B par $\widetilde{B}$; nous devrons alors remplacer $N^{-1} D$ par $N^{-1} \widetilde{D} = \widetilde{\Sigma}$, qui est bien définie positive. Nous avons alors :

$$(\log_e L)_{\widetilde{B},\widetilde{\Sigma}} \geqslant (\log_e L)_{\widetilde{B},\Sigma} \tag{126}$$

D'autre part, l'inégalité (121) peut s'écrire :

$$(\log_e L)_{\widetilde{B},\Sigma} \geqslant (\log_e L)_{B,\Sigma} \tag{127}$$

De (126) et (127), on déduit (120).

5 – Distribution d'échantillonnage de $\widetilde{\Sigma}$ – Epreuve de l'hypothèse (H).

Si l'hypothèse (H) est réalisée,

$$\mathcal{E}(\hat{B}K - M) = 0, \quad \text{ou,} \quad \mathcal{E}(\hat{B})K = M \tag{128}$$

Nous pouvons donc écrire dans ce cas :

$$\hat{B}K - M = \hat{B}K - \mathcal{E}(\hat{B})K = (\hat{B} - \mathcal{E}(\hat{B}))K \tag{129}$$

Puisque K est de rang s, T9 et T8 du chapitre 2 permettent de démontrer que $K' A^{-}K$ est une matrice régulière symétrique et définie positive. Il en va de même évidemment de $(K' A^{-}K)^{-1}$ et il existe donc une matrice régulière d'ordre s, soit S, telle que :

$$S(K' A^{-}K)^{-1} S' = I_s \tag{130}$$

Posons

$$Y = (y_{i\alpha}) = (\hat{B}K - M)S^{-1} = (\hat{B} - \mathcal{E}(\hat{B}))K S^{-1} \tag{131}$$

Nous désignerons par $\vec{y}_i$ la ligne de rang i et par $\vec{y}_\alpha$ la colonne de rang α de la matrice Y. Ce sont là des vecteurs ayant respectivement s et p composantes.

Si (H) est réalisée, on déduit de la relation (131) :

$$\mathcal{E}(Y) = \mathcal{E}(\hat{B} - \mathcal{E}(\hat{B}))\, K\, S^{-1} = 0 \tag{132}$$

En particulier, on a :

$$\mathcal{E}(\vec{y}_i) = 0 \; , \; (i = 1, \ldots, p) \tag{133}$$

$$\mathcal{E}(\vec{y}_\alpha) = 0 \; , \; (\alpha = 1, \ldots, s) \tag{134}$$

Calculons maintenant la covariance des lignes $\vec{y}_i$ et $\vec{y}_j$:

$$\begin{aligned} \mathcal{E}(\vec{y}_i{}'\vec{y}_j) &= (S^{-1})'\, K'\, \mathcal{E}\,\{(\vec{\hat{b}}_i - \mathcal{E}(\vec{\hat{b}}_i))'\,(\vec{\hat{b}}_j - \mathcal{E}(\vec{\hat{b}}_j))\}\, K\, S^{-1} \\ &= \sigma_{ij}\,(S^{-1})'\, K'\, A^- K\, S^{-1} = \sigma_{ij}\,(S\,(K'\, A^- K)^{-1} S')^{-1} \\ &= \sigma_{ij}\, I_s, \end{aligned} \tag{135}$$

comme le montrent successivement les formules (98) et (130).

La relation matricielle (135) entraîne les relations scalaires :

$$\mathcal{E}(y_{i\alpha}\, y_{j\beta}) = \sigma_{ij}\, \delta_{\alpha\beta} \; , \; (i\, , j = 1, \ldots, p\, ; \alpha\, , \beta = 1, \ldots, s) \tag{136}$$

De ces relations, on déduit aisément :

$$\mathcal{E}(\vec{y}_\alpha\, \vec{y}_\beta{}') = \delta_{\alpha\beta}\, \Sigma \tag{137}$$

Les formules (134) et (137) montrent que chaque colonne $\vec{y}_\alpha$ de Y est une variable aléatoire $N(0\, , \Sigma)$ et que $\vec{y}_\alpha$ et $\vec{y}_\beta$ sont indépendantes si $\alpha \neq \beta$. Finalement, compte tenu de (131).

$$(\hat{B}K - M)\,(K'\, A^- K)^{-1}\,(\hat{B}K - M)' = YS\,(K'\, A^- K)^{-1}\, S'Y' = Y\,Y' = \sum_{\alpha=1}^{s} \vec{y}_\alpha\, \vec{y}_\alpha{}' \tag{138}$$

Nous avons ainsi démontré :

T16 – Dans l'hypothèse (H), $(\hat{B}K - M)\,(K'\, A^-K)^{-1}\,(\hat{B}K - M)'$ obéit à une loi de WISHART $W\,(\Sigma\, , s)$.

Ce théorème étend le lemme 8.4.1. de T.W. ANDERSON (loc. cit., p. 191) au cas où $r = \text{rang}\; Z < q$.

Pour éprouver l'hypothèse (H), on comparera les maxima de la vraisemblance de l'échantillon calculés respectivement pour le modèle réduit, soit $L\,(\tilde{B}\, , \tilde{\Sigma})$, et pour le modèle sans restrictions, soit $L\,(\hat{B}\, , \hat{\Sigma})$.

En effectuant les subtitutions requises dans l'expression (30) de $\log_e L$, on obtient les résultats suivants :

$$(\log_e L)_{\tilde{B},\tilde{\Sigma}} = -\frac{1}{2} Np \log_e 2\pi - \frac{1}{2} pN - \frac{1}{2} N \log_e |\tilde{\Sigma}| \tag{139}$$

$$(\log_e L)_{\hat{B},\hat{\Sigma}} = -\frac{1}{2} Np \log_e 2\pi - \frac{1}{2} pN - \frac{1}{2} N \log_e |\hat{\Sigma}| \tag{139'}$$

On en déduit aisément :

$$\lambda = \frac{L(\widetilde{B}, \widetilde{\Sigma})}{L(\hat{B}, \hat{\Sigma})} = \frac{|\hat{\Sigma}|^{\frac{N}{2}}}{|\widetilde{\Sigma}|^{\frac{N}{2}}} = \frac{|\hat{D}|^{\frac{N}{2}}}{|\widetilde{D}|^{\frac{N}{2}}} \tag{140}$$

et enfin, en utilisant la formule (119) et en élevant les deux membres de (140) à la puissance $\frac{2}{N}$:

et $$\lambda^{\frac{2}{N}} = \frac{|\hat{D}|}{|\widetilde{D}|} = \frac{|\hat{D}|}{|\hat{D} + (\hat{B}K - M)(K' A^{-}K)^{-1}(\hat{B}K - M)'|} \tag{141}$$

L'étude de la distribution d'échantillonnage de $\lambda^{\frac{2}{N}}$ dans l'hypothèse (H) est fondée uniquement sur les théorèmes T13 et T16. Les considérations développées par T.W. ANDERSON (loc. cit. § 8.5. et 8.6.) lui sont directement applicables.

$\lambda^{\frac{2}{N}}$ est désigné plus explicitement par $U_{p,s,N-r}$, où p est la dimension des variables normales étudiées et où s et N − r sont les paramètres intervenant à côté de Σ dans la spécification des distributions de WISHART de T16 et T13.

Rappelons que lorsque p = 2, on a :

$$\frac{1 - \sqrt{U_{2,s,N-r}}}{\sqrt{U_{2,s,N-r}}} \times \frac{N - r - 1}{s} = F, \tag{142}$$

où F possède respectivement 2s et 2 (N − r − 1) degrés de liberté.

5 – APPLICATION

Nous allons appliquer notre méthode aux données de IMMER, HAYES et POWERS, citées par T.W. ANDERSON (loc. cit., p 217).

Il s'agit de N = 30 vecteurs observés à deux composantes qui représentent les rendements d'orge mesurés au cours de deux années consécutives. Ces vecteurs peuvent être disposés dans un tableau à double entrée, la première correspondant aux 6 lieux d'expérience et la seconde aux 5 variétés étudiées.

Le modèle théorique correspondant à ces observations peut se formuler comme suit :

$$\mathcal{E}(\vec{x}_{ij}) = \vec{\mu} + \vec{\lambda}_i + \vec{\gamma}_j, \ (i = 1, \dots, 6 \, ; j = 1, \dots, 5)$$

où les différents vecteurs aléatoires sont indépendants et ont même matrice des covariances. (On constate que $N = 30 > 2 + 12 = p + q$). On peut encore écrire :

$$\mathcal{E}(\vec{x}_{ij}) = B \vec{z}_{ij}, \ (i = 1, \dots, 6 \, ; j = 1, \dots, 5)$$

où

$$B = (\vec{\mu} \mid \vec{\lambda}_1 \mid \vec{\lambda}_2 \mid \vec{\lambda}_3 \mid \vec{\lambda}_4 \mid \vec{\lambda}_5 \mid \vec{\lambda}_6 \mid \vec{\gamma}_1 \mid \vec{\gamma}_2 \mid \vec{\gamma}_3 \mid \vec{\gamma}_4 \mid \vec{\gamma}_5)$$

est une matrice de format 2 x 12 et où $\vec{z}_{ij}$ est un vecteur colonne à 12 composantes qui sont toutes nulles à l'exception de celles de rang 1, 1 + i et 1 + i + j qui sont égales à 1.

Tableau des données
Variétés

		M	S	V	T	P	Totaux
Lieux	UF	81	105	120	110	94	514
		81	82	80	87	84	414
	W	147	142	151	192	146	778
		100	116	112	148	108	584
	M	82	77	78	131	90	458
		103	105	117	140	130	595
	C	120	121	124	141	125	631
		99	62	96	126	76	459
	GR	99	89	69	89	104	450
		66	50	97	62	80	355
	D	87	77	79	102	96	441
		68	67	67	92	94	388
Totaux		616	611	621	765	659	3 272
		517	482	569	655	572	2 795

Afin de passer à l'expression matricielle de ce modèle, il faut pouvoir ranger en une suite les vecteur $\vec{x}_{ij}$ dans une matrice X de format 2 x 30 et les vecteurs $\vec{z}_{ij}$ dans une matrice Z de format 12 x 30. Nous avons adopté la convention suivante :

$$\vec{x}_{ij} = \vec{x}_\alpha \quad \text{et} \quad \vec{z}_{ij} = \vec{z}_\alpha, \quad \text{avec} \quad \alpha = 5\,(i - 1) + j$$

Compte tenu de cette convention et du tableau des données, on calcule aisément les matrices $A = ZZ'$ et $C = XZ'$ dont les expressions détaillées figurent dans l'annexe 1. Dans cette annexe, on trouvera également les expressions du g-inverse A^- choisi, de AA^- et de $\hat{B} = CA^-$.

Pour obtenir A^-, nous avons utilisé la méthode exposée par S.R. SEARLE dans le chapitre 6 de "Matrix Algebra for the Biological Sciences". Cette méthode, fondée sur la réduction de A à la forme diagonale, a permis de vérifier que rang A = r = 10. Ainsi la matrice A d'ordre 12 est bien singulière. Du tableau des données et de l'annexe 1, on obtient :

$$XX' = \begin{pmatrix} 380.944 & 315.381 \\ 315.381 & 277.625 \end{pmatrix}$$

et

$$\hat{B}A\hat{B}' = \hat{B}C' = \begin{pmatrix} 377.665 & 314.579 \\ 314.579 & 273.608 \end{pmatrix}$$

De la formule (73), on déduit donc :

$$N\,\hat{\Sigma} = XX' - \hat{B}A\hat{B}' = \begin{pmatrix} 3\,279 & 802 \\ 802 & 4\,017 \end{pmatrix}$$

Jusqu'ici nous avons étudié le modèle général sans restrictions. Nous allons maintenant formuler des hypothèses au sujet des paramètres du modèle.

Dans le tableau de l'annexe 2, nous avons étudié quelques matrices K formées d'une seule colonne. La comparaison de la première et de la deuxième partie de ce tableau, c'est-à-dire de K et de $AA^{-}K$, permet de voir si la condition (8), soit $AA^{-}K = K$, est vérifiée ou non. On constate qu'elle est vérifiée pour les 6 dernières colonnes du tableau. Ainsi, parmi les fonctions BK que nous avons considérées, seules celles correspondant à ces 6 dernières colonnes sont estimables et pour elles seules nous avons calculé $\hat{B}K$.

Il est évident que les matrices K obtenues en juxtaposant des colonnes vérifiant la condition (8), vérifient elles mêmes cette condition. Il en va ainsi notamment de la matrice K obtenue en juxtaposant les 4 dernières colonnes.

On vérifie aisément que ces 4 colonnes sont linéairement indépendantes et par suite que s = rang K = 4. En outre :

$$BK = (\vec{\gamma}_1 - \vec{\gamma}_2 \mid \vec{\gamma}_1 - \vec{\gamma}_3 \mid \vec{\gamma}_1 - \vec{\gamma}_4 \mid \vec{\gamma}_1 - \vec{\gamma}_5)$$

Si nous voulons éprouver l'hypothèse qu'il n'y a pas de différences entre les variétés, soit :

$$(H) : \vec{\gamma}_1 = \vec{\gamma}_2 = \vec{\gamma}_3 = \vec{\gamma}_4 = \vec{\gamma}_5,$$

nous utiliserons alors une matrice M = 0 de format 2 x 4.

Puisque M = 0, l'expression (119) de $N\,\tilde{\Sigma}$ devient :

$$N\,\tilde{\Sigma} = N\,\hat{\Sigma} + \hat{B}K\,(K'\,A^{-}K)^{-1}\,K'\,\hat{B}'$$

L'estimation $\hat{B}K$ s'obtient évidemment en juxtaposant les 4 dernières colonnes d'estimations de l'annexe 2. Par ailleurs, la matrice K ayant un rang égal à 4, T9 garantit que $K'\,A^{-}K$ est régulière. On trouvera son expression détaillée ainsi que celle de son inverse dans cette même annexe. On peut dès lors calculer :

$$\hat{B}K\,(K'\,A^{-}K)^{-1}\,K'\,\hat{B}' = \begin{pmatrix} 2\,788 & 2\,550 \\ 2\,550 & 2\,863 \end{pmatrix},$$

et enfin :

$$N\,\tilde{\Sigma} = \begin{pmatrix} 6\,067 & 3\,352 \\ 3\,352 & 6\,880 \end{pmatrix}$$

Pour éprouver l'hypothèse (H), il reste à former le quotient :

$$\lambda^{\frac{2}{N}} = \frac{|\hat{D}|}{|\widetilde{D}|} = \frac{|N\hat{\Sigma}|}{|N\widetilde{\Sigma}|} = 0{,}4107$$

Puisque p = 2 , s = 4 , N = 30 et r = 10, on a $\lambda^{\frac{2}{N}} = U_{2,4,20}$ et compte tenu de (142), on calculera

$$\frac{1 - \sqrt{U_{2,4,20}}}{\sqrt{U_{2,4,20}}} \times \frac{19}{40} = 2{,}66$$

Ce résultat dépasse 2,20, valeur de $F_{8,38}$ correspondant au seuil de 0,05. L'hypothèse (H) est donc rejetée ; il y a des différences significatives entre les variétés d'orge. Nous obtenons ainsi un résultat identique à celui fourni par l'analyse classique d'ANDERSON.

Notre but en développant cette application était simplement d'illustrer la méthode.

D'un point de vue théorique, on peut constater qu'elle permet de découvrir d'emblée les fonctions paramétriques BK qui sont estimables.

D'un point de vue pratique, la démarche est la même quel que soit le plan d'expérience étudié. C'est là une circonstance favorable à la programmation sur ordinateur.

BIBLIOGRAPHIE

ANDERSON T.W. – An Introduction to Multivariate Statistical Analysis. Wiley, New York, 1958.

RAO C.R. – Linear Statistical Inference and its Applications. Wiley, New York, 1965.

RAO C.R. and MITRA S.K. – Generalized Inverse of Matrices and its Applications. Wiley, New York, 1971.

SEARLE S.R. – Matrix Algebra for the Biological Sciences. Wiley, New York, 1966.

SEARLE S.R. – Linear Models. Wiley, New York, 1971.

ANNEXE 1 – LES MATRICES A , A^- , AA^- , C et $\hat{B}$

$$A = \begin{pmatrix} a & b & b & b & b & b & b & c & c & c & c & c \\ b & b & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ b & 0 & b & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ b & 0 & 0 & b & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ b & 0 & 0 & 0 & b & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ b & 0 & 0 & 0 & 0 & b & 0 & 1 & 1 & 1 & 1 & 1 \\ b & 0 & 0 & 0 & 0 & 0 & b & 1 & 1 & 1 & 1 & 1 \\ c & 1 & 1 & 1 & 1 & 1 & 1 & c & 0 & 0 & 0 & 0 \\ c & 1 & 1 & 1 & 1 & 1 & 1 & 0 & c & 0 & 0 & 0 \\ c & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & c & 0 & 0 \\ c & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & c & 0 \\ c & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & c \end{pmatrix} \quad A^- = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & d & e & e & e & e & e & f & f & f & f & 0 \\ 0 & e & d & e & e & e & e & f & f & f & f & 0 \\ 0 & e & e & d & e & e & e & f & f & f & f & 0 \\ 0 & e & e & e & d & e & e & f & f & f & f & 0 \\ 0 & e & e & e & e & d & e & f & f & f & f & 0 \\ 0 & e & e & e & e & e & d & f & f & f & f & 0 \\ 0 & f & f & f & f & f & f & d & g & g & g & 0 \\ 0 & f & f & f & f & f & f & g & d & g & g & 0 \\ 0 & f & f & f & f & f & f & g & g & d & g & 0 \\ 0 & f & f & f & f & f & f & g & g & g & d & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

avec a = 30, b = 5 et c = 6 $\qquad$ avec $d = \frac{1}{3}$, $e = \frac{4}{30}$, $f = -\frac{1}{6}$, $g = \frac{1}{6}$

$$AA^- = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & h & h & h & h & 0 \end{pmatrix}$$

avec h = – 1

$$C = \begin{pmatrix} 3272 & 514 & 778 & 458 & 631 & 450 & 441 & 616 & 611 & 621 & 765 & 659 \\ 2795 & 414 & 584 & 595 & 459 & 355 & 388 & 517 & 482 & 569 & 655 & 572 \end{pmatrix}$$

$$\hat{B} = \frac{1}{30} \begin{pmatrix} 0 & 3107 & 4691 & 2771 & 3809 & 2723 & 2669 & -215 & -240 & -190 & 530 & 0 \\ 0 & 2549 & 3569 & 3635 & 2819 & 2195 & 2393 & -275 & -450 & -15 & 415 & 0 \end{pmatrix}$$

ANNEXE 2 – LES MATRICES K

K	1	0	0	1	1	0	0	0	0
	0	1	0	1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	1	1	1
	0	0	0	0	0	– 1	0	0	0
	0	0	0	0	0	0	– 1	0	0
	0	0	0	0	0	0	0	– 1	0
	0	0	1	0	1	0	0	0	– 1
AA^-K	0	1	0	1	0	0	0	0	0
	0	1	0	1	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	1	1	1
	0	0	0	0	0	– 1	0	0	0
	0	0	0	0	0	0	– 1	0	0
	0	0	0	0	0	0	0	– 1	0
	0	1	0	1	0	0	0	0	– 1
BK	$\vec{\mu}$	$\vec{\lambda}_1$	$\vec{\gamma}_5$	$\vec{\mu} + \vec{\lambda}_1$	$\vec{\mu} + \vec{\gamma}_5$	$\vec{\gamma}_1 - \vec{\gamma}_2$	$\vec{\gamma}_1 - \vec{\gamma}_3$	$\vec{\gamma}_1 - \vec{\gamma}_4$	$\vec{\gamma}_1 - \vec{\gamma}_5$
$\hat{B}K$	x	x	x	x	x	$\frac{25}{30}$	$-\frac{25}{30}$	$-\frac{745}{30}$	$-\frac{215}{30}$
	x	x	x	x	x	$\frac{175}{30}$	$-\frac{260}{30}$	$-\frac{690}{30}$	$-\frac{275}{30}$

Pour la matrice K formée des 4 dernières colonnes du tableau :

$$K' A^- K = \frac{1}{6} \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} \quad \text{et} \quad (K' A^- K)^{-1} = \frac{6}{5} \begin{pmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{pmatrix}$$